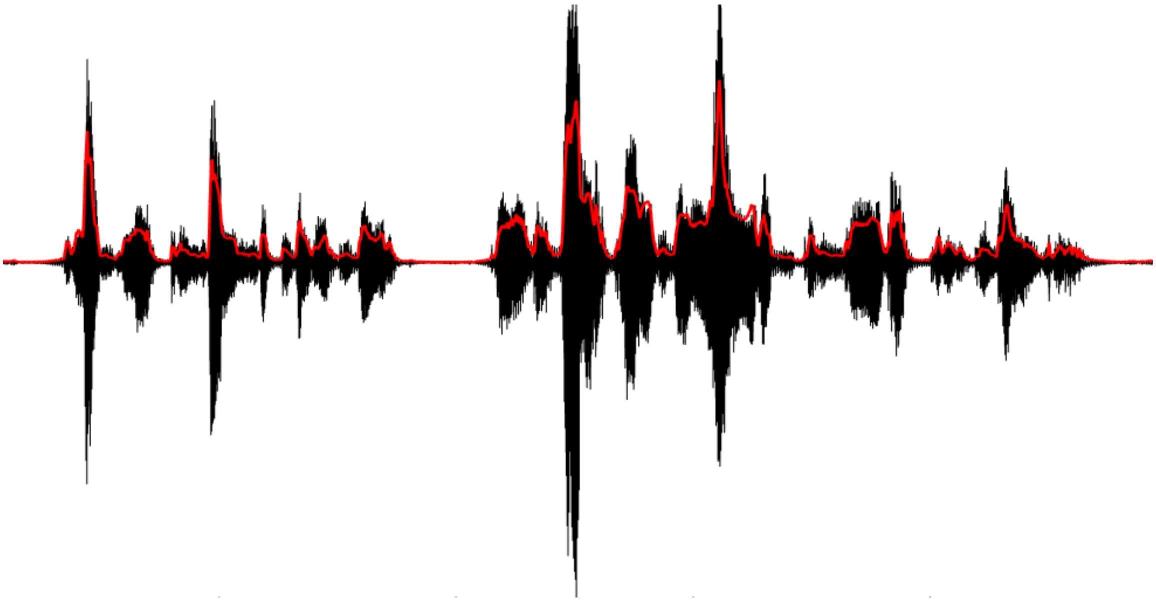
Rhythmic Sensitivity and Developmental Language Disorder in Children

Susan Richards

Clare Hall

July 2017



This dissertation is submitted for the degree of Doctor of Philosophy

THESIS SUMMARY

Rhythmic Sensitivity and Developmental Language Disorder in Children

Susan Richards

Children with Developmental Language Disorder (DLD) have difficulties in acquiring language in the absence of other neurodevelopmental issues (e.g. autism, hearing impairment) and despite growing up in an adequate language-learning environment. Previous characterisations of DLD have focused on grammatical processing, phonological memory or rapid auditory processing. This thesis approaches the language-learning difficulties of children with DLD from a novel perspective by considering the potential contribution made by differing levels of sensitivity to the rhythmic properties of language.

Children with DLD have been shown to have reduced sensitivity to some of the acoustic cues present in speech which are thought to be important for rhythmic perception. Since rhythm forms the basis of language processing in early development, poorer sensitivity to language rhythm may result in later language problems.

To investigate whether children with DLD demonstrate difficulties in processing language rhythm, this thesis explores five areas of language processing which could be affected by poor rhythmic sensitivity: locating word-boundaries, processing novel words, storing lexical stress patterns, representing sentence level structures and the integration of rhythm and syntax. As part of the investigation, measures were also taken of acoustic threshold sensitivity to see whether task performance related to acoustic sensitivity. A parallel strand of the study investigated whether provision of an entraining rhythm prior to task stimuli could support task performance.

Three groups of children participated in the study: children with DLD, age-matched TD children (AMC) and younger, language-matched TD children (YLC). The results indicate that rhythmic manipulation of language stimuli affects task responses across the five language areas under investigation. The findings are then discussed in terms of the contribution made to our understanding of the role of rhythm in language and language disorder.

ACKNOWLEDGEMENTS

I dedicate this thesis to my husband, Gavin and to my children, Katie and Alexander.

My utmost gratitude to my supervisor, Usha Goswami, for her guidance, insight, encouragement, and most of all, for always believing in me.

She is an inspiration.

I would also like to thank all the members of the Centre for Neuroscience in Education, past and present, for making it such a vibrant, stimulating and supportive place to be.

My thanks also to Ann-marie Cano CertMRCSLT, who carried out the second transcriptions, and to the ESRC, who generously funded my studies at Cambridge.

Lastly, I would like to thank all the teachers, parents and children who made this research possible.

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text

This dissertation does not exceed 60,000 words exclusive of figures, tables, footnotes, bibliography and appendices.

TABLE OF CONTENTS

Thesis Overview

This study approaches the language-learning difficulties of children with DLD from a novel perspective by considering the potential contribution made by differing levels of sensitivity to the rhythmic properties of language.

In Chapter One, we will begin by considering the nature of rhythm in language. We will then turn to a consideration of what acoustic and neural accounts of speech processing tell us about the underlying nature of rhythmic processing and how a rhythmic processing difficulty might arise. Finally, we will consider the implications of rhythmic processing difficulties in the context of developing language.

Chapter Two forms a general introduction to the main study, outlining the participant characteristics and procedural details which are common to all of the experiments outline in Chapters Four to Nine.

Chapter Three gives some theoretical background to the concept of Entrainment, which is a common feature throughout the experimental tasks.

Chapter Four outlines the four tasks which were used to estimate Acoustic Thresholds for the participants.

Chapters Five to Nine each deal with a different experimental task. The theoretical basis for each task is introduced, followed by an explanation of how the specific task materials were derived and constructed. The results for each task are then presented together with a discussion of the findings.

Chapter Ten forms a general discussion of the themes which have arisen throughout the course of the experimental tasks regarding the role of rhythm in language and language disorder.

1 Literature Review

1.1 Conceptualising Linguistic Rhythm

Since rhythm is the primary focus of this thesis, we will begin with a discussion of what may constitute rhythm in the context of language.

The notion of linguistic rhythm, whilst widely acknowledged as part of each speaker's language competence, has proved surprisingly difficult to define, with extensive debate across the literature of both linguistics and psychology about how to quantify speakers' intuitions. Some approaches have evaluated levels of isochronicity between stress and syllable intervals (Dauer, 1983; Wenk & Wioland, 1982), whilst others have attempted in different ways to quantify the relative durational values of consonants and vowels to describe language rhythm (Ramus, Nespor, & Mehler, 1999; Grabe & Low, 2002). These attempts have primarily focused on describing the ways in which rhythmic characteristics vary between languages, rather than on the nature of rhythm itself. An alternative approach, and the one which this thesis draws upon, focuses on rhythm as a hierarchical construct, consisting of patterns of elements unfolding over time.

1.1.1 Rhythm as Patterning

Rhythm is a property of many aspects of human endeavour outside of language processing and we can look to these alternatives to help frame an understanding of rhythm in language.

In music, the existence of rhythmic structure is uncontroversial, and Lerdahl & Jackendoff (1983) provided a seminal discussion of musical rhythmic structure. They proposed a notion of rhythm as a patterned series of strong and weak beats, organised not just as a temporal sequence, but as a hierarchy in which weaker beats are nested within stronger beats. This idea can be illustrated using a simple example (Figure 1-1).

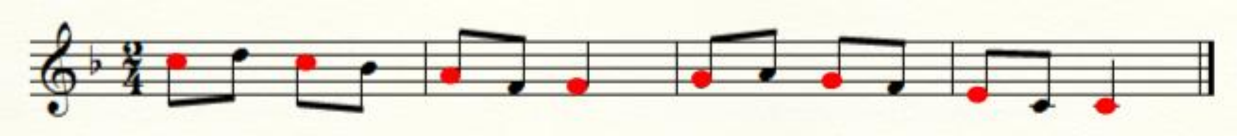


Figure 1-1 Musical notation of the nursery rhyme 'Polly put the kettle on'. Red notes indicate the location of strong beats.

Strong beats (red notes) occur at regular intervals, whilst differing numbers of weak beats occur between the strong beats. It is this nesting of weak beats between the strong beats which creates the musical rhythm.

This conceptualising of rhythm as strong and weak beats has parallels with how we can think about linguistic rhythm. Language is also characterised by alternations of stressed or strong syllables (analogous to strong beats) and unstressed or weak syllables (analogous to weak beats). The complementary nature of these rhythmic structures is revealed when we consider how language is set to music. In constructing songs, stressed syllables tend to align with strong musical beats, and unstressed syllables with weaker musical beats. We can see this if we add the lyrics to the tune above:

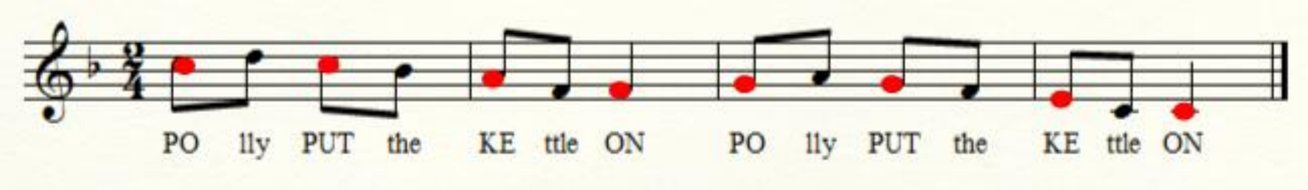


Figure 1-2 Musical notation of the nursery rhyme 'Polly put the kettle on' with lyrics. Strong beats are indicated in red; Strong syllables are indicated by capital letters.

The stressed syllables (PO, PUT, KE, ON) align with the strong beats, whilst the unstressed syllables (lly, the, ttle) align with the weak beats. We can therefore see immediately two layers of structure operating – individual notes which correspond to individual syllables, and above them, strong beats corresponding to stressed syllables.

A parallel notion of linguistic rhythm is found in the field of poetry. Poetic rhythm or metre is “the ordered patterning of stressed and unstressed syllables in verse” (OED, 2001). Stressed and unstressed syllables in poetry are combined to form a ‘metrical foot’ which may be trochaic (Sw), iambic (wS) or more complex such as the anapaest (wwS) or dactyl (Sww). Considering syllable stress patterns as the basis for linguistic rhythm in English is therefore a well-established part of poetic form - “a syllable which is perceived as stressed may be perceived as the beat of a metrical pattern, and hence of an underlying rhythm” (Attridge, 1982). There appears a clear understanding in poetry that stressed syllables are perceived as rhythmic beats and frame the metrical structure of the verse.

These complementary views of rhythm from musical and linguistic arts both consider rhythm as a hierarchical structure centred around the concept of strong and weak beats, with strong beats providing a framework within which the weaker beats are nested. Rhythm is therefore a hierarchical *patterning* of events rather than a linearly quantified series of temporal durations. The focus on patterning therefore also dispenses with the quest for isochronicity. Whilst speech rhythms are not isochronous, they are nonetheless regular and predictable, enabling listeners to create expectations of when linguistic events will occur (Peelle & Davis, 2012), and a parallel can be drawn with music in which expressive variations in timing are anticipated by listeners and incorporated into their expectancies of musical structure (Repp, 1992; Clarke, 1989).

Arvaniti (2009) drew on the use of rhythm in music to propose her characterisation of language rhythm, describing rhythm as ‘the product of prominence and patterning’ (p61) with rhythm in English relying on the ‘grouping and (…) alternation of more and less prominent syllables’ (p58). Crucially, she therefore argued that rhythm was about *patterning* rather than *timing* per se.

In this thesis, we will consider linguistic rhythm as consisting of a patterning of temporally spaced acoustic events grouped within a hierarchical framework. In this conceptualisation, the strong or stressed syllables of speech form the higher-order structure and nested between the stressed syllables are varying numbers of weak or unstressed syllables. The alternation of these strong and weak syllables creates the rhythmic patterns of a word, phrase or utterance. In this sense, rhythm is seen as the “systematic temporal, accentual and grouping pattern(s) of sound” (Patel, 2008, p150), primarily realised as the patterning of stressed and unstressed syllables.

The following sections will introduce Developmental Language Disorder (DLD) and consider what is currently known about rhythm in relation to children with DLD before discussing possible locations of a rhythmic difficulty in speech and neural processing systems. Finally the potential impact of a rhythmic difficulty in the context of language development will be discussed.

1.2 Rhythm and Developmental Language Disorder

1.2.1 Developmental Language Disorder

Children with DLD have difficulties in acquiring language in the absence of other neurodevelopmental issues (e.g. autism, hearing impairment) and despite growing up in an adequate language-learning environment. Recent estimates suggest prevalence of 7% (Dockrell, Lindsay, Palikara, & Cullen, 2007; Norbury et al., 2016; Tomblin et al., 1997).

Children with DLD may have difficulties across a range of linguistic areas such as word-learning (Kan & Windsor, 2010) grammatical morphology (Marchman, Wulfeck & Ellis Weismer, 1999; van der Lely, Rosen, & Adlard, 2004) and comprehension of syntax (Bishop, Bright, James, Bishop, & van der Lely, 2000). Children with DLD face a variety of challenges in accessing education and employment, as well as being at risk of emotional and mental health difficulties (Conti-Ramsden, Knox, Botting, & Simkin, 2002; Law, Rush, Schoon & Parsons, 2009).

There is currently a lack of consensus on the causes of DLD, although many perceptual and cognitive factors have been implicated in language disorder, such as rapid auditory processing (Tallal & Piercy, 1973), phonological memory (Gathercole & Baddeley, 1990) and grammatical deficits (van der Lely, Rosen, & McClelland, 1998). One aspect of language competence which has thus far escaped attention is the potential role played by sensitivity to rhythm in language disorder. Given the importance of aspects of rhythm in speech processing (outlined in Section 1.3) and language development (outlined in Section 1.4), we propose that there is considerable theoretical justification for exploring implications of rhythmic sensitivity in children experiencing language difficulties.

1.2.2 Non-linguistic Rhythm and DLD

Studies of rhythm in children with DLD are not plentiful, however some evidence for a difficulty with non-linguistic rhythmic processing has been found. The KE family (widely studied for the strongly hereditary form of DLD displayed in some family members) participated in several tests of pitch and rhythm. Affected family members performed more poorly on tests of rhythmic perception and production than did controls, indicating a level of rhythmic difficulty for those family members who also displayed language difficulties (Alcock, Passingham, Watkins, & Vargha-Khadem, 2000). In an entrainment task, Corriveau & Goswami (2009) asked children with DLD to tap to a metronome beat, and found that they were considerably poorer at synchronising their taps with the metronome than either age-matched or language-matched control children at rates of 2Hz and 1.5Hz. These slower rates broadly correspond to the inter-stress intervals typically found in speech (Dauer, 1983). Beat synchronisation skills were also found to contribute unique variance to measures of language and literacy. The finding of poor motor synchronisation with the beat in children with DLD, even for richer, musical stimuli, was replicated in Cumming, Wilson, Leong, Colling, & Goswami (2015).

Richards (2010) asked children with DLD to perform a music-based rhythm task. The children were first played a recording of a short nine-note rhythm, played at a single pitch using a synthesised piano timbre. They then heard two rhythms, one the original, one a foil which differed from the original in the placement of one strong-weak note pairing, with each rhythm ‘played’ by a different cartoon character on screen. The original rhythm was then repeated and the children had to indicate which character had played that rhythm. Children with DLD were significantly poorer at carrying out this task than age-matched controls (*p* = .001), performing marginally above chance, whilst typically-developing (TD) children approached ceiling. A similar speech-based rhythm task, using low-pass filtered sentences (thus removing phonetic information) and a flat f0 contour did not yield a significant group difference, although Cumming et al. (2015) did find a group difference on an adapted version of the speech-based task with a larger cohort of children.

Further speech and music-based rhythm tasks performed by Cumming et al. (2015) found group differences in tasks of musical beat perception and of tapping to the beat of music, with DLD children performing less well than TD peers. Furthermore, they found that speech rhythm matching and musical beat perception were significant predictors of scores in receptive and expressive subtests of the CELF, with children with DLD who had better rhythm matching or musical beat perception having better language outcomes. Better rhythm and pitch matching was also associated with better outcomes in tests of phonological awareness. Weinert (1992) found that children with DLD who did more poorly in a rhythm discrimination task were also poorer at learning an artificial language.

Further evidence for links between rhythmic processing and language skills has been found in typically-developing children. Gordon et al. (2015) found that performance in a test of rhythm discrimination correlated significantly with scores in expressive morpho-syntax in typically developing 6-year-old children, accounting for 48% of variance in scores, lending further support to the notion that proficiency in rhythmic processing may be linked to better language outcomes. The inverse inference from the data provided by Gordon et al. and Cumming et al. is therefore that children with poorer rhythmic processing will have poorer language outcomes.

There is therefore evidence that children with DLD have difficulty in differentiating between musical rhythms, in entraining tapping to a metronome beat and in performing other rhythmic tasks.

1.2.3 Language Rhythm in DLD

There has to date been little interest in exploring language rhythm in relation to language disorder and so it is currently unknown whether the difficulties found with non-linguistic rhythm are also found in tasks investigating rhythmic patterns in language. One exception is the study carried out by Richards & Goswami (2015).

In this study we investigated the representations of rhythm at the word-level (i.e. lexical stress) in children with DLD. We showed children a picture of an item (e.g. a potato) and then played two different realisations of the target word. One token had the correct stress pattern (i.e. poTAto, wSw) whilst the other had an incorrect stress pattern (e.g. potaTO, wwS)[1]. The children had to decide which of the two tokens was correct. We found that the children with DLD scored significantly more poorly on this task than typically-developing children, suggesting that they had less robust representations of lexical stress.

In that work we also investigated the relationship between lexical stress representations and discrimination of rise time – one of the major acoustic cues to stress (see Section 1.3.1.2 for a detailed explanation of rise time). The children carried out a rise time discrimination task in addition to the experimental lexical stress task, and we found that rise time performance was a significant predictor of lexical stress performance for our DLD group, explaining 37.3% of unique variance. This is the first evidence to suggest that poorer acoustic sensitivity (here to rise time) may be related to linguistic aspects of rhythm (here patterns of lexical stress).

There is some limited evidence, then, that children with DLD have difficulty with language rhythm as exemplified by lexical stress patterns and that this may be related to reduced sensitivity to some of the acoustic cues that contribute towards stress and rhythm perception in language.

Building on this preliminary result, this thesis describes an attempt to explore the wider ramifications of acoustic sensitivity, rhythm and language processing and the potential implications for language disorder.

Section 1.3 will provide an overview of the acoustic and neural processes of speech which are relevant for rhythmic perception together with how these may be affected in children with DLD. Furthermore, since the children have a *Developmental* Language Disorder, we need to consider the

---

[1] Throughout this thesis, the stress status of syllables will be indicated by S (Strong) and w (weak) so that, for example, Sw indicates a strong-weak syllable pairing (e.g. TAble), and wS a weak-strong syllable pairing (e.g. beLIEF). When syllables are written in full, capitals denote a strong syllable.

role that language rhythm has in the development of language and what impact a reduced level of rhythmic sensitivity may have on that developmental process. This will be outlined in Section 1.4.

1.3 Rhythm and the Acoustic and Neural Processing of Speech

We have defined rhythm as the patterning in time of stressed and unstressed syllables, and so it is apposite to consider how rhythm in speech relates to acoustic and neural accounts of speech processing to understand what systems may be implicated in reduced sensitivity to rhythm. This section will explore how rhythmic processing is currently understood in relation to speech and how this may be affected in children with DLD.

1.3.1 Rhythm and Acoustic Aspects of Speech

Speech is a complex acoustic signal, containing a wealth of information across frequency, amplitude and temporal domains. Each of these aspects has its part to play in accounting for the perception of rhythm.

1.3.1.1 The Role of Frequency

Traditional accounts of speech have focused on the frequency information typically conveyed by spectrograms (Figure 1-3).

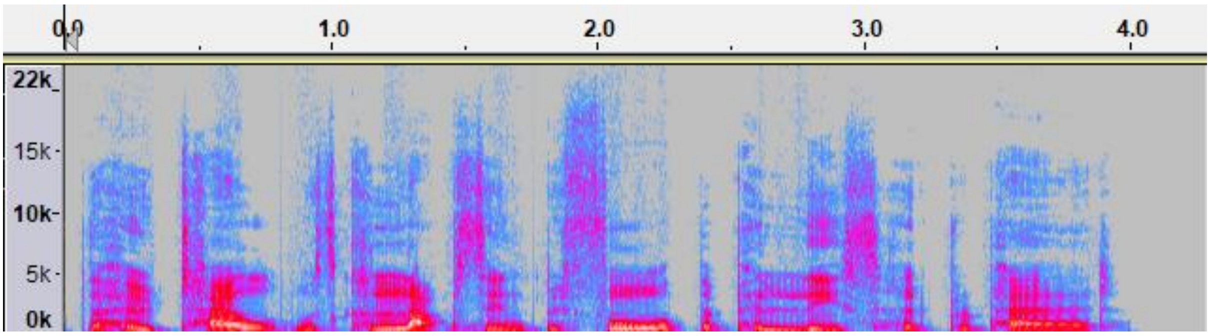


Figure 1-3 Spectogram of the sentence "The hotel was collapsing before the police had prepared"

Spectograms provide information about the level of energy present across a wide range of frequencies (e.g. up to 22kHz), and have enabled researchers to focus on aspects of the signal such as the rapid transient features observed at the release of consonants, and on the formant characteristics of different vowels. In this respect, acoustic processing of speech has been regarded as occurring at the level of individual phonemes. Since rhythm occurs across larger units, this traditional mode of acoustic analysis is less useful for studying its properties.

However, fundamental frequency does have a role to play in rhythm as it is one of the acoustic characteristics which contributes to the perception of stress. Stressed syllables tend to be higher in pitch than unstressed syllables (Laver, 1994) and early investigations into stress (e.g. Fry (1958)) found that dynamic increases in fundamental frequency were the most significant contributors to stress perception in wordpairs such as ‘OBject’ and ‘obJECT’, above duration and intensity. More recent studies have argued for a lesser significance for frequency compared with other acoustic characteristics (e.g. Greenberg, Carvey, Hitchcock, & Chang, 2003; Kochanski, Grabe, Coleman & Rosner, 2005) however a contribution to the overall percept remains.

Difficulty in discriminating changes in frequency could therefore lead to difficulties in distinguishing levels of stress, and by extension processing the rhythms that stress patterns create. Primary school children with DLD have been found to have poorer frequency discrimination skills than age-matched peers (Cumming, Wilson, & Goswami, 2015; Mengler, Hogben, Michie, & Bishop, 2005), whilst older children with DLD tended to not differ on tests of discrimination (McArthur & Bishop, 2004). In our previous study (Richards & Goswami, 2015) we did not find a frequency discrimination difference between children with DLD and age-matched controls, whilst Cumming, Wilson, & Goswami (2015) found that only the subgroup of DLD children who had additional phonological processing deficits had poorer frequency discrimination thresholds. The presence of difficulties with frequency discrimination in DLD therefore appears to be dependent on the characteristics of particular sample populations. For children who do have poor frequency discrimination, this could be a contributing factor to any difficulties with stress and rhythm perception.

1.3.1.2 The Role of Amplitude

Spectrograms provide detailed information about the energy present at different frequencies in the speech signal, however a complementary way of displaying speech signal information uses the sound pressure wave.

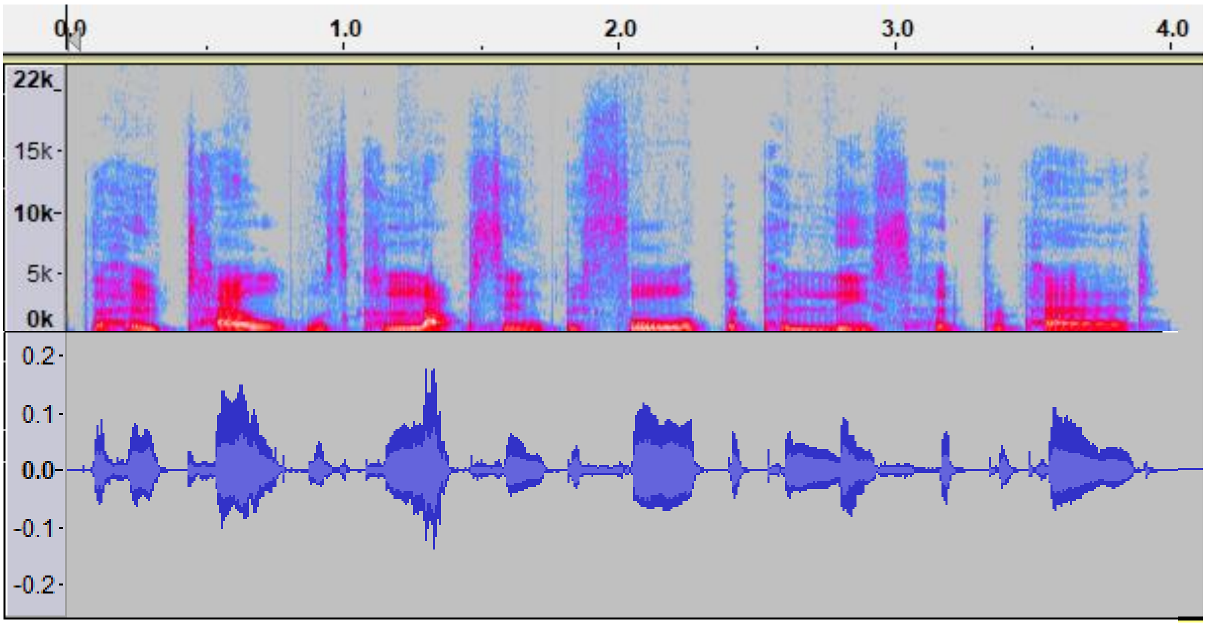


Figure 1-4 Spectogram and Waveform of the sentence "The hotel was collapsing before the police had prepared"

Figure 1-4 shows the spectrogram (above) and the waveform (below) of the same sentence. Whilst the spectrogram displays frequency information, the waveform displays the sound pressure levels with the overall shape driven by changes in amplitude over time.

A further diagram of the same sentence (Figure 1-5) highlights the shape of these amplitude changes (red line), which is known as the ‘amplitude envelope’. The amplitude envelope represents the acoustic power present in the signal (usually summed across all frequencies) and is a more slowly varying aspect of the signal, with dominant modulations typically in the 4-8 Hz (125 – 250ms) range. This timescale is therefore broadly equivalent to the timescale of the syllable in typical speech (Poeppel, 2003).

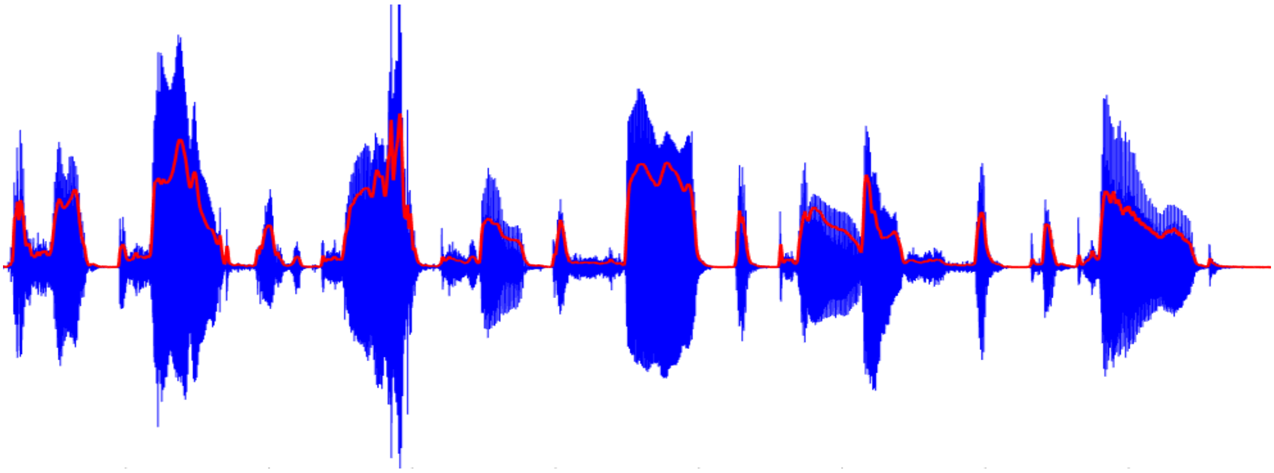


Figure 1-5 Waveform of the sentence ‘the hotel was collapsing before the police had prepared’ with the amplitude envelope marked in red

Until relatively recently, the role of the amplitude envelope in speech had received little attention, however a steady body of work over the last 20 years has demonstrated that modulations of the amplitude envelope in fact play a crucial role in language processing, and significantly for this thesis, language rhythm at the level of the syllable and stressed syllable (Rosen, 1992).

1.3.1.2.1 The role of the amplitude envelope in detail

The fine structure (rapidly varying modulations- typically 600Hz – 10kHz) and the amplitude envelope (slowly varying modulations – typically 2Hz-50Hz) are complementary aspects of the signal, with each contributing different information regarding segmental and suprasegmental aspects of speech (Rosen, 1992). The fine structure (see Figure 1-6 below) contains significant segmental information for individual phonemes such as place and voicing cues. In a complementary fashion, the envelope, with fluctuations between 2 and 50Hz, generally provides suprasegmental information relevant for syllabic and suprasyllabic prosodic features (although some segmental information on manner and vowel identity can be transmitted). Because of its longer timescale operating at the level of the syllable and beyond, it is the envelope which provides information on rhythm (Rosen, 1992).

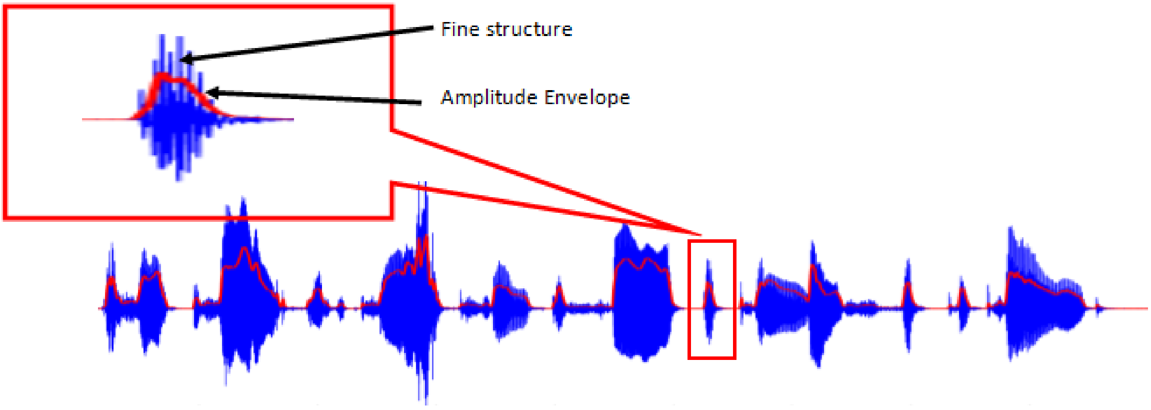


Figure 1-6 Waveform for sentence "the hotel was collapsing before the police had prepared" highlighting the word 'the' to illustrate the amplitude envelope and the fine structure

Deconstructing the speech signal into these two complementary strands of information (fine structure and envelope) has yielded insights into speech processing systems. Smith, Delgutte, & Oxenham (2002) asked participants to listen to various auditory chimera created by combining the amplitude envelope of one utterance with the fine structure of another. They found that participants reported the information contained in the amplitude envelope, suggesting that it was the envelope which was driving speech recognition. Shannon, Zeng, Kamath, Wygonski, & Ekelid (1995) retained the amplitude envelope of speech whilst degrading the spectral cues. They found that the temporal cues present in the amplitude envelope were sufficient to produce 90% correct identification of target words. In a complementary set of studies, low-frequency temporal information was removed from the signal and this was found to lead to a significant decrease in intelligibility in speech (Drullman, Festen, & Plomp, 1994; Ghitza, 2012). The information contained in the amplitude envelope therefore appears to be making significant contributions to speech recognition.

Modulations of the amplitude envelope occurring at the syllabic-rate therefore have a crucial role to play in speech processing. Changes in the amplitude envelope demarcate syllables, enabling the processing of syllable-level information. If rhythm occurs as the successive patterning of weak and strong syllables, then the capacity to process speech at the syllabic level would appear fundamental to this process.

The amplitude envelope is not just relevant for discovering syllable-level speech units. It also has a central role to play in distinguishing between levels of stress – the patterning of which is the foundation of our notion of rhythm. We saw above that fundamental frequency contributes to the percept of stress, however a more significant indicator of stress levels (strong or weak) is the amplitude rise time. The amplitude rise time represents the time-frame of the slope between the onset of the amplitude rise at the beginning of the syllable and its peak at the syllable nucleus. Stressed syllables have shorter rise times, with a rapid change in amplitude leading to a high intensity peak at the syllable nucleus. Unstressed syllables have slower changes in amplitude, leading to a lesser peak of intensity at their nucleus (Greenberg, Carvey, Hitchcock, & Chang, 2003).

Figure 1-7 illustrates this property of syllable rise time for the bisyllabic Sw word ‘SEven’.

Graph removed as copyright clearance has not been granted. Copyright holder Steven Greenberg.

Figure 1-7 Schematic representation of the acoustic properties of the word ‘seven’. Reproduced from Greenberg, Carvey, Hitchcock & Chang, 2003, p 480, Fig. 14

From the figure, the high amplitude peak with its sharp increase in amplitude (i.e. a short rise time) can be seen for the stressed syllable ‘SE’ with a lower peak and gentler slope (i.e. a longer rise time) for the unstressed syllable ‘ven’.

The capacity to distinguish between different rise times is therefore central to the ability to distinguish between differing levels of stress. A reduced sensitivity to differences in rise times, conversely, is likely to lead to impoverished representations of stress and therefore of rhythm in the wider sense.

1.3.1.2.2 Amplitude rise time and DLD

An extensive range of studies have linked poor discrimination of rise time to developmental dyslexia (e.g. Goswami, 2011; Goswami et al., 2002; Goswami, Gerson, & Astruc, 2010). In their initial paper, Goswami et al. (2002) speculated that an even more significant deficit in rise time than found in children with dyslexia could cause the kind of spoken language difficulties found in DLD.

There is substantial overlap in groups of children with dyslexia and those identified with DLD. Around half of children diagnosed with DLD may also have reading difficulties (McArthur, Hogben, Edwards, Heath, & Mengler, 2000;Tomblin, Zhang, Buckwalter, & Catts, 2000) and approximately half of children diagnosed as having dyslexia may also have oral language difficulties (McArthur et al. 2000). Longitudinally, children who have preschool language difficulties persisting until the beginning of literacy instruction frequently go on to present with reading difficulties (Snowling, Duff, Nash, & Hulme, 2015). As heritable disorders, having a family risk of dyslexia also leads to an increased risk of DLD with both traits being characterised by a higher incidence of weak preschool phonological skills (Nash, Hulme, Gooch, & Snowling, 2013).

Given the degree of overlap between the disorders, it therefore seems reasonable to investigate whether a similar underlying difficulty in processing rise time cues as identified in children with dyslexia, but taking a different developmental course, could play a role in the language problems of children with DLD. Research in this area has been sparse, however a few studies have investigated rise time sensitivity with these children.

Corriveau, Pasquini, & Goswami (2007) used two different rise time tasks with children with DLD. One investigated discrimination of single rise times (one-ramp task) and a second investigated discrimination of two successive rise times (two-ramp task). They found that the children with DLD had higher thresholds for both tasks, as did Beattie & Manis (2013), whilst Fraser, Goswami, & Conti-Ramsden (2010) found group differences between DLD and age-matched controls (AMC) on the two-ramp task only. Group differences between DLD and AMC groups were also found for speech-based rise time tasks used by Cumming, Wilson & Goswami (2015). In our own previous work (Richards & Goswami, 2015) we also found that children with DLD had poorer sensitivity to rise time than AMC children.

There is therefore some evidence that children with language difficulties are less able to discriminate between differences in rise time than TD children, and that poorer rise time sensitivity may relate to stress perception (Richards & Goswami, 2015).

1.3.1.3 The role of the temporal aspects of speech

A third acoustic factor related to stress perception is duration. Stressed syllables are longer than unstressed syllables and this differentiation contributes to perception of stress levels (Greenberg et al., 2002). Studies investigating duration discrimination thresholds have found children with DLD to have less sensitive discrimination of durational differences than TD children (Corriveau, Pasquini & Goswami, 2007; Richards & Goswami, 2015; Cumming, Wilson & Goswami, 2015). Difficulty in differentiating between acoustic events of different durations could also therefore hamper perception of stress in children with DLD.

Temporal considerations of rhythm extend beyond the duration of individual units, however. One of the fundamentals of rhythm is that it is a pattern which unfolds over time – a single acoustic event cannot be said to have rhythm. Indeed, Laver defines rhythm as the “interaction in *time* of the relative prominence of stressed and unstressed syllables” (Laver, 1994, p152).

Timing in speech is complex, operating across multiple timescales, from individual phonemes at 20-40ms, syllables at 150-250ms up to the longer units of words, phrases and clauses (Poeppel, 2003). Despite the varying timescales involved, these differing units have to be integrated into a coherent percept for subsequent processing. Traditional accounts have considered this as a sequential ‘smallest up’ process in which phonemes are combined to create syllables, syllables combined to form words and so on.

In contrast, recent models of speech processing have incorporated the differing temporal aspects of speech into a ‘multi-time resolution’ system in which speech is analysed at short windows (20-40ms, corresponding to segmental-level units) and simultaneously also analysed at longer windows (150-250ms, corresponding to syllable-level units), (Poeppel, Idsardi, & van Wassenhove, 2008). Within this system, it can be supposed that as a suprasegmental property of speech, rhythm would arise from processes at the longer analysis windows at the syllable-level and above.

The timing of syllable-level information is also an important factor in speech intelligibility. Ghitza & Greenberg (2009) inserted a range of silent intervals into time-compressed sentences. They found that the intelligibility of the sentences varied according to the timeframe (and therefore ensuing rhythm) of the silent intervals. They concluded that it was not merely the *nature* of the acoustic signal which mattered for speech comprehension, but also the *timing* of that signal’s presentation.

Tracking of speech in time therefore appears to be a central process in efficient speech comprehension. Furthermore, the concept of temporal processing occurring simultaneously across multiple timescales is relevant for our notion of rhythm as nesting of stressed and unstressed syllables. If we take as an example the nursery rhyme ‘Mary Mary’ (Figure 1-8), there are two ‘rates’ at which salient rhythmic events occur – the syllable rate (typically 250ms intervals) and, simultaneously, the stressed syllable rate (typically 500ms intervals).

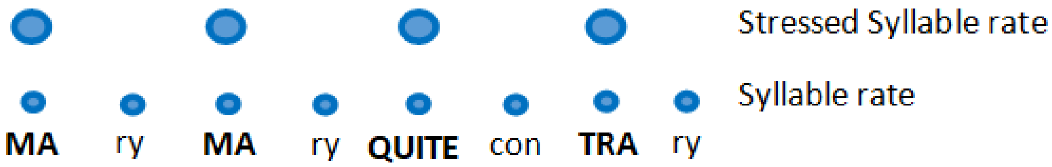


Figure 1-8 Illustration of 'Mary Mary'. Small circles denote syllables, large circles denote stressed syllables

Leong (2012) demonstrated that it is the nesting or *phase* of the syllable and stressed syllable rates (i.e. the degree to which the two levels work together) which gives rise to the percept of a particular rhythm, such as that found in nursery rhymes. Significantly, in her work, it was specifically the relevant timing of the amplitude envelope peaks at the two levels which created the temporal structure for the rhythms.

Efficient processing of rhythm therefore appears to depend upon efficient processing of amplitude rise times and the capacity to track and integrate the temporal distribution of amplitude rise times across multiple timescales. Poor discrimination and integration of these differing rates of temporal information is therefore likely to lead to poorer rhythmic perception. There is no data on this phenomenon for children with DLD, however adults with dyslexia are poorer at using multi-scale temporal information to detect rhythms (Leong & Goswami, 2014).

The necessity of accurately tracking rise times leads us to a further aspect of processing which is also relevant to the processing of rhythm – that of neural oscillations.

1.3.2 Rhythm and Neural oscillations

Neural oscillations are physiological properties of neural systems and represent rhythmic fluctuations in the excitability of neuronal populations. Oscillations are cyclical and represent the alternation between states of low and high neuronal excitability. Recent research has demonstrated that oscillations respond to the characteristics of external stimuli (such as auditory speech) at multiple time-scales and that when oscillations are in time (or *phase*) with external stimuli, neural processing is more efficient.

Neural oscillations cycle through phases of high and low excitability, however the temporal location of these cycles can be reset by an external stimulus (such as an auditory rhythm) so that the high excitability phase of the oscillation coincides with the timeframe of an acoustic event in a process known as ‘phase-resetting’ (Lakatos et al., 2005; Schroeder & Lakatos, 2009; Thut, Miniussi, & Gross, 2012). Once the phase of oscillations has been reset to the rhythm of an external stimulus, the oscillations ‘phase-lock’ to this until further input causes them to reset. The process of phase-resetting and phase-locking therefore enables neural systems to both react dynamically to the input being received, and, because the process is cyclical, thereafter anticipate the timeframe of the next input to maximise responsiveness.

The capacity of neural systems to phase-lock or ‘entrain’ to the temporal spacing of external events (visual or auditory) is thought to facilitate the accurate and timely processing of those events. In the visual system of macaque monkeys, once oscillations are entrained to the rhythmic presentation of stimuli (demonstrated by amplified neuronal responses) reaction times to target stimuli vary according to when in the phase of oscillation the stimuli occur (Lakatos, Karmos, Mehta, Ulbert, & Schroeder, 2008). Schroeder & Lakatos (2009) suggest that the entrainment of high excitability states is a means of directing sensory selection and attention. In this way, attention will be greatest at the timepoint directed by the rhythmic structure of a stimulus, coinciding with the time of highest neuronal excitability, thereby maximising processing efficiency. In contrast, stimuli arriving at the timepoint of a trough in the oscillatory cycle, will be met with a lower level of neuronal responsiveness. By maximising system responsiveness to timepoints of maximal information input, attention operating in ‘rhythmic mode’ makes the most efficient use of neural resources (Schroeder & Lakatos, 2009).

A further important discovery for understanding how oscillations may relate to speech processing is that oscillations occur at multiple timescales. These timescales are often referred to as different bands of frequencies: *gamma* (the most rapidly fluctuating signals at 30-50Hz), *theta* (4-10Hz) and *delta* (1-4Hz) (Lakatos et al., 2005; Gross et al., 2013). This range of response bands allows the brain to simultaneously process complex rhythms across different timescales by providing complementary information about a stimulus (Panzeri, Brunel, Logothetis, & Kayser, 2010). From Poeppel’s multi-time resolution model of speech processing (2003), we can see that the different bands of oscillatory activity correspond to the timeframes associated with different linguistic units: phonemes at the gamma rate (30-50Hz), syllables at the theta rate (4-10Hz), and stressed syllables at the delta rate

(1-4Hz) (Ghitza, 2011; Ghitza, Giraud, & Poeppel, 2013; Ghitza & Greenberg, 2009). We can therefore conceive of a multi-level system which entrains (through the process of phase resetting) to rhythmic fluctuations in the speech signal occurring at phoneme, syllable and stress levels (Poeppel, 2003; Power, Mead, Barnes, & Goswami, 2012).

Phase-tracking of stimuli corresponding to the gamma and theta rates of speech has been demonstrated in adults (Luo & Poeppel, 2012) and theta-band oscillations have been shown to discriminate between sentence stimuli, with poorer phase tracking associated with poorer speech intelligibility (Luo & Poeppel, 2007), suggesting a syllable-level segmentation of sentences and strengthening the notion of the syllable as a fundamental unit for speech perception (Greenberg et al., 2003). Phase-locking of auditory cortex and speech temporal envelopes is also correlated with speech comprehension (Ahissar et al., 2001) and intelligibility (Doelling, Arnal, Ghitza, & Poeppel, 2014) leading to suggestions that phase-locking of cortical responses to the temporal structure of speech may be a prerequisite of speech comprehension (Ahissar et al., 2001). Ghitza (2012) showed that intelligibility of speech stimuli was poor when band envelopes in the theta range were flattened, but that when the input-rhythm information was restored, intelligibility improved. He interpreted this as demonstrating the importance of theta-rate tracking for speech comprehension in the role of syllabic parsing. In further studies (e.g. Doelling et al., 2014), the temporal cues of amplitude envelope rise times at the syllabic rate were thought to be the instigators of envelope-tracking, enabling the parsing of speech stimuli into chunked representations at the level of the syllable. Furthermore, the degree of neural envelope tracking shown by participants correlates with ratings of intelligibility (Doelling et al., 2014). It should be noted that oscillatory tracking is not specific to speech – musical sequences also cause entrainment of cortical oscillations to the dominant note rate (Doelling & Poeppel, 2015).

Oscillations are thought to be reset by the onset of a sound, sometimes referred to as an “auditory edge” (Luo & Poeppel, 2012). Rapid changes in the amplitude envelope are though to constitute such an auditory edge or "acoustic landmark" (Doelling et al., 2014), acting as the agent which entrains the phase of neural oscillations. It has been suggested that insensitivity to auditory edges (i.e. amplitude rise times) may result in impaired phase-locking and thus poorer perception of speech (Doelling et al., 2014). If children with DLD have poorer sensitivity to rise times, then this could lead to a reduced phase-locking of neural oscillations, thereby impacting on their speech perception. This in turn could lead to difficulties across language systems. In a musical tracking study, successful tracking of the temporal structure of music resulted in the facilitation of an apparently unrelated skill (pitch perception) (Doelling & Poeppel, 2015). It is thus plausible to predict a similar effect for language, in which the ability to track the temporal structure of speech, thus predicting timing of significant speech events (stress and syllables) facilitates semantic and syntactic processing. Conversely, a diminished capacity to track temporal structures of rhythm is likely to lead to poorer processing of language.

For example, in a paper based on neurophysiological data, Kotz, Schwartze, & Schmidt-Kassow, (2009) directly related rhythmic, temporal expectancy to syntactic processing. They suggest that language processing involves two forms of prediction: *what* next? and *when* next?. Syntactic prediction (*what* next?) is based on knowledge of typical language structures (e.g. the article 'the' is likely to be followed by a noun to form a noun phrase), which interfaces with temporal prediction (i.e. *when* next?; (Large & Kolen, 1994), with the idea being that if *what?* and *when?* coincide, the greatest efficiency of processing will occur.

If this theory accurately captures aspects of language and attentional processing, then we can speculate how these factors may interact over the course of language development. Children who are better able to predict *when?* will be better placed to discover *what?* and vice versa, in a cyclical process of expectancy reinforcement. In contrast, children whose systems are less able to abstract temporal regularities to predict *when?* will be at a disadvantage in discovering the *what?*. We hypothesise that this could be the case for children with DLD, where reduced sensitivity to the cues of language rhythm are likely to result in less accurate induction of temporal and attentional expectancy (*when?*) and thus their language systems will be less primed to efficiently discover the *what?* of language structure. This could therefore offer a temporal rhythm explanation for the syntactic difficulties experienced by children with DLD (e.g. van der Lely, Rosen, & McClelland, 1998).

1.3.3 Summary

Acoustic processing of stress and rhythm depends on sensitivity to changes in frequency, amplitude and duration together with the capacity to track these changes simultaneously across multiple timescales. Children with DLD have previously been found to have reduced sensitivity to these acoustic features, with rise time in particular being related to poorer performance in judgements of lexical stress.

Poor sensitivity to acoustic features such as rise time may also have wider processing implications through the contribution of neural oscillations to language processing. Neural oscillations phase-lock to auditory stimuli through auditory edges such as the amplitude rise time. Effective phase-locking can facilitate language processing through effective direction of neural resources to salient points in time. Poor induction of rhythmic expectancies through these acoustic and neural mechanisms could lead to poorer processing across a range of language skill areas.

We will now turn to a consideration of the role of rhythm in language development and the potential implications of poor rhythmic sensitivity for this process.

1.4 Rhythm and Language Development

"In the Beginning, was Rhythm" *Hans von Bülow*

Humans are sensitive to sound from early in development with foetuses responding to sound from the third trimester. This includes linguistic material such as mother's voice, or familiar stories and rhymes both in the womb, and shortly after birth (DeCasper & Fifer, 1980; DeCasper & Spence, 1986; Decasper, Lecanuet, Busnel, Granier-Deferre & Maugeais, 1994). What form this sound takes remains somewhat speculative, however it seems that the foetal environment effectively acts as a low pass filter, with high frequency sounds progressively attenuated (Cooper & Aslin, 1990). It therefore seems likely that in terms of language, it is the prosodic patterning of rhythm and intonation contour which would survive such a filtering effect, forming the first linguistic input to the unborn child. Processing of prosodic structure would therefore be the initial foundation upon which the subsequent language system is built. This concept of prosody as the foundation of language makes the study of prosodic structure of particular interest in considering what might cause that acquisition process to go awry, as in the case of children with DLD. If prosody is the bedrock on which language is built, and yet a child's perception of prosody is weak or inefficient, the resulting language system may also be inefficient.

1.4.1 Rhythm at Birth

Since newborn infants respond differently to stimuli familiar or unfamiliar from time spent in the womb, the operation of processing and recording linguistic stimuli has evidently begun before birth. Newborn infants have therefore already spent several weeks listening to and processing the prosodic patterns of their native language. This exposure means that newborns can already discriminate between their native language and a foreign language (Mehler, Jusczyk, Lambertz, & Halsted, 1988). This is a prosodic, rather than segmental, judgement since discrimination occurs even when the language has been low-pass filtered, effectively removing all phonetic content, leaving only rhythm and contour. Nazzi, Bertoncini, & Mehler (1998) found that languages with similar rhythms (e.g. English, Dutch) were indistinguishable from each other, but that languages with differing rhythmic patterns (e.g. English, Italian) were readily discriminated, indicating that it was rhythmic characteristics alone which were prompting discrimination. Infants are therefore sensitive to the rhythmic properties of language from birth, and before.

This primacy of rhythm in newborn infants' processing of language prompted Mehler, Dupoux, Nazzi & Dehaene-Lambertz (1996) to state that 'the basic representation of speech is based on the rhythmic structures (…) embodied in the utterances of a language' (p108). Rhythm is therefore a central component of the very earliest representations of language.

1.4.2 Rhythm and Discovering Word Boundaries

A particular challenge in language acquisition is discovering and segmenting individual units of language (such as words) from the continuous speech stream. A major theory of language acquisition proposes that the rhythmic properties of language can be used to demarcate the speech stream into smaller, meaningful units, and can be used to discover consistencies at the word level.

This proposal is based on the rhythmic properties of English words characterised by their stress patterns. Cutler & Carter (1987) investigated the rhythmic (stress distribution) properties of English words and found large asymmetries between the location of stressed and unstressed syllables at word level. The most frequent word structure in English is a polysyllable with initial primary stress, and the most frequent polysyllable is a bisyllabic word with the primary stress on the first syllable, followed by a second, weak, syllable (Carlson, Elenius, Granström & Hunnicott, 1985). Almost 90% of content words begin with a strong syllable, as do approximately 90% of all lexical tokens in spontaneous conversation. In contrast, across all word-types, weak initial syllables are found in less than 27% of words in English, whilst weak initial polysyllabic words have a frequency of only 6 per million words in written frequency counts (Cutler & Carter, 1987).

This asymmetry of distribution with strong syllables disproportionately occurring at word-initial locations led Cutler to propose the 'Metrical Segmentation Strategy'. This strategy proposes that learners assume that any strong syllable in the input must be word-initial, and use this assumption to segment the speech stream into word-like units.

Much of Cutler's work to support this hypothesis has been completed with adult speakers of languages across different rhythmic groups (English-stress, French-syllable, Japanese-mora[2]). She showed that listeners tend to exploit the rhythmic pattern of their native language (whatever that happens to be) in order to segment efficiently. If English speakers use stress patterns, then French speakers rely on syllables (Cutler, Mehler, Norris, & Segui, 1986), and Japanese speakers use morae (Cutler & Otake, 1994). Furthermore, speakers attempt to impose their native language strategies when listening to languages from other rhythmic classes, for example French speakers attempt to segment English syllabically (Cutler et al., 1986) and Japanese speakers attempt to segment English moraically (Cutler & Otake, 1994). Proficient speakers are therefore attentive to the rhythmic

---

[2] A 'mora' is a sub-syllabic unit consisting of a vowel (V), consonant (C) or CV combination and is considered the basic rhythmic unit in Japanese.

properties of language and impose their existing templates to interpret any linguistic content to which they are exposed.

These observations have several important implications. First of all, they serve to highlight a primary role for rhythmic segmentation in language, since speakers across the globe appear to be using rhythm as a core part of their linguistic knowledge. Secondly, since the manifestation of this appears to be language-specific, this must be something that early language-learners (i.e. infants) can derive from the specific input that they experience. Thirdly, since adults attempt to use that knowledge when deciphering unfamiliar languages, rhythmic structure must become a core part of that person's linguistic expectations rather than being derived anew based on current input.

Rhythmic segmentation has also been shown to be a particularly robust strategy in the face of limited phonetic input. Briscoe (1989) used machine learning algorithms in an attempt to discover lexical items in conditions of phonetic uncertainty. In these circumstances, the constraints imposed by the metrical segmentation strategy based on strong syllables led to more successful results than those based on phoneme, word or syllable-level analysis. Harrington, Watson, & Cooper (1989) also found that an algorithm based on strong vowel analysis (as a proxy for stress) was more successful in detecting word boundaries than one based on phonotactic constraints.

'Slips of the ear' also have a rhythmic basis, with adult listeners tending to incorrectly insert boundaries prior to strong syllables (Cutler & Butterfield, 1992). Cutler & Butterfield discuss this as particularly relevant when the input is difficult to interpret – just as in the machine-learning experiments, in the situation where there is some kind of degradation of the signal, a rhythmic strategy appears to be an effective means of compensating for deficits in the input. They acknowledge that such a strategy does not produce flawless results, but rather argue that it operates as a reasonable holistic approximation, a 'rule of thumb', which allows for rapid and efficient processing for the majority of the input.

Infant studies have since corroborated this adult and machine-learning data. 7.5-month-old infants employ a Sw parsing strategy, regarding strong syllables as word onsets (Jusczyk, Houston, & Newsome, 1999). This strategy led to the infants 'misparsing' sequences such as 'guitar is' as containing a word 'taris', based on the strong-weak sequencing. By 10.5 months, however, infants are wise to the existence of exceptions and were able to extract weak-strong words successfully. English-speaking 7.5 month olds can segment on the basis of Sw preference, whilst French-speaking 7.5-month-olds do not, supporting rhythmic segmentation as a language-specific phenomenon (Polka, Sundara & Blue, 2002). By 7.5 months, therefore, infants have already developed expectancies about the distribution of rhythmic structures within their native language.

Word segmentation based on rhythmic properties therefore appears to be a consistent part of speakers' linguistic competence. Speakers create templates and expectancies about rhythmic cues at the word-level based on exposure to their native language which they then use in order to demarcate and process all types of linguistic input. If infants are poor at processing this fundamental property of language, it follows that they are likely to fail to establish these core expectations of segmentation which could serve to facilitate their further language acquisition. Infants who are poor segmenters have been found to have poorer word production at 12 and 24 months (Weber, Hahne, Friedrich, & Friederici, 2005) and lower language scores in early childhood (Newman, Bernstein Ratner, Jusczyk, Jusczyk, & Dow, 2006). Segmentation in infancy (which appears dependent on rhythm) may therefore predict later language development (Benasich & Tallal, 1996).

Poor sensitivity to rhythm could therefore be detrimental to the process of word segmentation, with subsequent effects on general language development. Based on data from English-French bilingual adults who used only one of their available strategies (either stress (English) or syllable (French)) for segmenting both of their 'native' languages, regardless of the target language (Cutler, Mehler, Norris, & Segui, 1992), Cutler hypothesised that infants need 'only one (rhythmic) starting point' (Cutler, 1996, p94) which is the one they continue to use throughout life. If children with DLD do not acquire this starting point as infants, or acquire a distorted one, it seems possible that they will therefore either fail to use it, or fail to use it effectively, throughout their lifespan. They would thereby be deprived of one fundamental source of language information which TD infants are able to use to bootstrap language development. An investigation of whether older children with DLD are able to distinguish word boundaries based on rhythmic cues could provide information on whether this is a potential area of difficulty in language disorder.

1.4.3 Rhythm and Word Acquisition

If Cutler's Metrical Segmentation Strategy is correct, and can be applied to strategies of language acquisition, then infants must perceive and attend to the stress patterns of words. Indeed, stress patterning appears to be a central component of early language awareness, with responsiveness to these cues present from birth.

Newborn infants can discriminate between disyllables with either strong-weak or weak-strong stress patterns and recognise wordlists as having matching stress patterns, even if the consonantal detail of the syllables varies (Sansavini, Bertoncini, & Giovanelli, 1997), possibly due to the typical presence of reduced vowels in weak syllables (van Ooijen, Bertoncini, Sansavini, & Mehler, 1997). This indicates that at the newborn stage, infants are sensitive to patterns of stress and can discriminate categorically based on stress patterns alone.

Morgan & Saffran (1995) investigated the responses of 6- and 9-month-olds to syllable and rhythm sequences related to word-level representations. Infants were played sequences of sequentially recurring syllables across different rhythmic patterns as well as non-sequential sequences. Six-month-olds appeared to use only the rhythmic regularities to create units, disregarding any sequential, phonological information, whilst 9-month-olds grouped units only when there was regularity in both rhythmic and segmental properties – sequentially identical sequences occurring across contradictory rhythmic patterns did not create a word percept. They therefore argued that for the older infants it was the convergence of these regularities that created their wordlike percepts. It therefore seems that rhythmic information is used as a primary source for grouping input in the earlier stages of acquisition, with sequential, phonological information incorporated at a later timeframe. For these older infants, maintenance of the grouping percept was most robust when rhythmic and sequential grouping converged. This suggests that correlated regularity of rhythm and sequence is most beneficial for maintaining and consolidating wordlike representations in infants. If rhythm is the skeleton on which phonological detail is grafted, then a difficulty with processing word-level rhythm could lead to problems in processing the phonological characteristics of words.

Word-level (lexical) stress patterns are part of stored as well as on-line lexical knowledge. We have seen that Sw patterns predominate in English, and English-speaking infants develop a preference for words conforming to this pattern. 6-month-olds show no preference for particular word patterns in sucking studies, but by 9 months prefer strong-weak experimental wordlists, even when the words themselves are low-pass filtered (Jusczyk, Cutler, & Redanz, 1993). Infants are therefore developing expectancies about the individual stress patterns of words. Stress patterning also forms part of infants' internalised lexical representation for individual words, with infants discriminating between familiarised words and segmentally identical items presented with a differing stress pattern (Curtin, Mintz, & Christiansen, 2005). Poorer sensitivity to word-level rhythms could therefore result in impoverished representations of a word's characteristic rhythmic pattern.

Sensitivity to word-level patterns of stress seems therefore to be a key part of the infant's toolbox in the early stages of learning language. Infants are sensitive to stress patterning from birth and are able to use rhythmic regularities to create wordlike groupings by the age of six months. As their system refines, expectancies are created such that they prefer input which corresponds to the predominant rhythmic patterning by nine months and are now able to combine rhythmic and sequential properties to create unified representations of words. If infants have poorer sensitivity to patterns of lexical stress, then they are less likely to be able to graft phonological detail on to stress patterns, resulting in underspecified lexical processing at a phonological level. This could therefore manifest as a reduced ability to accurately process novel words. Furthermore, as words are acquired, their associated stress patterns may be under-specified, leading to potential difficulties with lexical storage and retrieval. Investigation into these areas could provide information on whether the processing of stress at a lexical level is an area of difficulty for children with DLD.

1.4.4 Rhythm and Larger Syntactic Structures

Beyond the word-level, rhythmic structures may also facilitate the processing of larger linguistic units such as phrases, clauses and sentences.

Whilst prosodic and syntactic structures do not necessarily coincide (Selkirk, 1996), there is evidence that adults (Price, Ostendorf, Shattuck-Hufnagel, & Fong, 1991) and infants make use of prosodic cues to help group sentences into smaller units. Infants prefer stimuli in which pauses occur clause-finally (rather than mid-clause), with the same preference demonstrated for phrase-final pauses (Hirsh-Pasek et al., 1987; Jusczyk et al., 1992) indicating that by 7-10 months-old, infants are sensitive to the typical coincidence of prosodic and syntactic cues. Jusczyk et al. describe this process as a ‘perceptual precategorisation’ (p.287) which enables a more detailed analysis of each perceptual group. By segmenting perceptual groups in line with meaningful grammatical units, this precategorisation would serve to delimit alternatives, effectively chunking the continuous stream so that more nuanced, grammatical or statistical analysis can take place. This process is, of course, not faultless – Juszcyk, Houston & Newsome (1999) show us that misparsings on this basis can readily occur (such as ‘guitar is’ as ‘taris’), however it is proposed that it serves as a preliminary ‘rough and ready’ means of constraining the possibilities, enabling faster, more efficient processing. Morgan & Demuth (1996) supported this notion by stating that prosody ‘may contain clues (…) providing additional constraints on syntactic and semantic analyses (…) helping to ensure that these analyses get started in the proper direction’ (p2).

The notion of prosodic cues operating across sentences has two major implications for children with DLD. If rhythmic structures operate as a parameter constraint, enabling more efficient processing of the chunked language, then poorer sensitivity to rhythm could result in greater difficulty in representing and processing linguistic units above the word-level. Poorer chunking, through the perceptual absence of rhythmic cues, could result in language systems attempting to represent and structure whole sentences, rather than smaller, more manageable units. This is likely to lead to less efficient and less accurate sentence-level representations. Investigating whether children with DLD respond to rhythmic variation at a sentence-level could yield insights into their representation of rhythm at this larger-pattern level.

Furthermore, if rhythmic and prosodic cues are integrated with syntactic structures, then poorer sensitivity to rhythm may result in problems extracting regularities of syntax. This is of particular interest, since DLD is characterised by receptive and expressive difficulties with grammar. An investigation of whether children with DLD are indeed struggling to integrate rhythmic and syntactic information could deliver insights into whether rhythm processing at this more global level could be impacting on their grammatical development.

1.5 Implications of Rhythm and Language Development for Language Disorder

We have seen that rhythm plays a pivotal role in the early stages of language development. From this discussion we have isolated five specific areas of language processing which could be adversely affected by a difficulty in processing rhythm:

1) ***Discovering word boundaries*.** If the Metrical Segmentation Hypothesis is correct, then children who have difficulty in identifying the rhythmic patterns of their native language are less likely to use this strategy in segmenting individual words from the speech stream, thus removing a potential source of bootstrapping for early word identification.
2) ***Processing novel words***. Infants are sensitive to the stress patterns of words, integrating rhythmic and segmental information to create word-like percepts. Difficulty in processing different stress patterns could result in difficulty representing the rhythmic and segmental information which forms part of novel word processing.
3) ***Storing Lexical Stress patterns****.* As new words are learnt, the characteristic stress pattern forms part of that word’s representation. If stress patterns are not processed effectively, then it seems likely that this will have a deleterious effect on the storage of word rhythms as part of lexical representations.
4) ***Representing rhythm at a sentence level****.* Rhythm may act as a parameter constraint, facilitating the chunking of sentences into smaller, meaningful units which can be processed more efficiently. Poorer sensitivity to rhythm may therefore lead to difficulties in representing larger units of language such as clauses and sentences.
5) ***Integrating Rhythm and Syntax****.* Infants use prosodic cues, including rhythm, to assist in the task of demarcating grammatical units such as phrases and clauses. A difficulty in perceiving these rhythmic cues may lead to developing less well-established rhythmic-syntax relationships, preventing rhythmic chunking and thus making the discovery of syntax more challenging.

We chose to concentrate on each of these five areas to investigate the nature of rhythmic processing of language in DLD, devising a series of experimental tasks to probe each of these aspects of language processing in relation to rhythm.

1.6 Summary and Overview

Language rhythm rests on the successive patterning of stressed and unstressed syllables at different hierarchical levels. Children with DLD may have reduced sensitivity to amplitude envelope rise times which, together with frequency and duration cues, are key contributors to the accurate perception of syllable-level structures in speech and to the differentiation of stress levels. Furthermore it is thought that rise times function to reset neural oscillations, thereby maximising neural processing efficiency. Relative insensitivity to rise time is therefore likely to lead to difficulties in processing the rhythmic patterns of language across multiple timescales, resulting in less efficient language processing. Integrating individual auditory events (e.g. syllables) into grouped patterns (i.e. rhythms) may also function to direct attentional resources through creating rhythmic expectancies. A difficulty integrating syllables into rhythmic patterns may also lead to poorer language outcomes.

Developmentally we have seen that rhythm plays a fundamental role in early language acquisition in areas such as discovering word boundaries, representing new words, storing lexical patterns and interpreting wider linguistic structures at phrasal, clausal and sentence level. If infants have difficulties in processing the acoustic cues to rhythm (such as rise time), then it seems plausible that they will have difficulty in drawing upon aspects of language rhythm in developing their language systems, with adverse effects on later language skills.

In order to investigate this proposed relationship between rhythmic processing and language disorder, we chose five areas of language processing in which the developmental literature suggests rhythm plays a core role.

The five areas are:

1) Discovering word boundaries
2) Processing stress patterns of novel words
3) Storing Lexical Stress patterns
4) Representing Rhythm at a Sentence Level
5) Integrating Rhythm and Syntax

An experimental task was devised for each language area, designed to probe the impact of language rhythm on task success.

In addition, we used Acoustic Threshold Estimation tasks (AT) to measure sensitivity to four acoustic properties: Frequency, Rise Time, Intensity and Duration, to explore whether acoustic sensitivity was related to task performance.

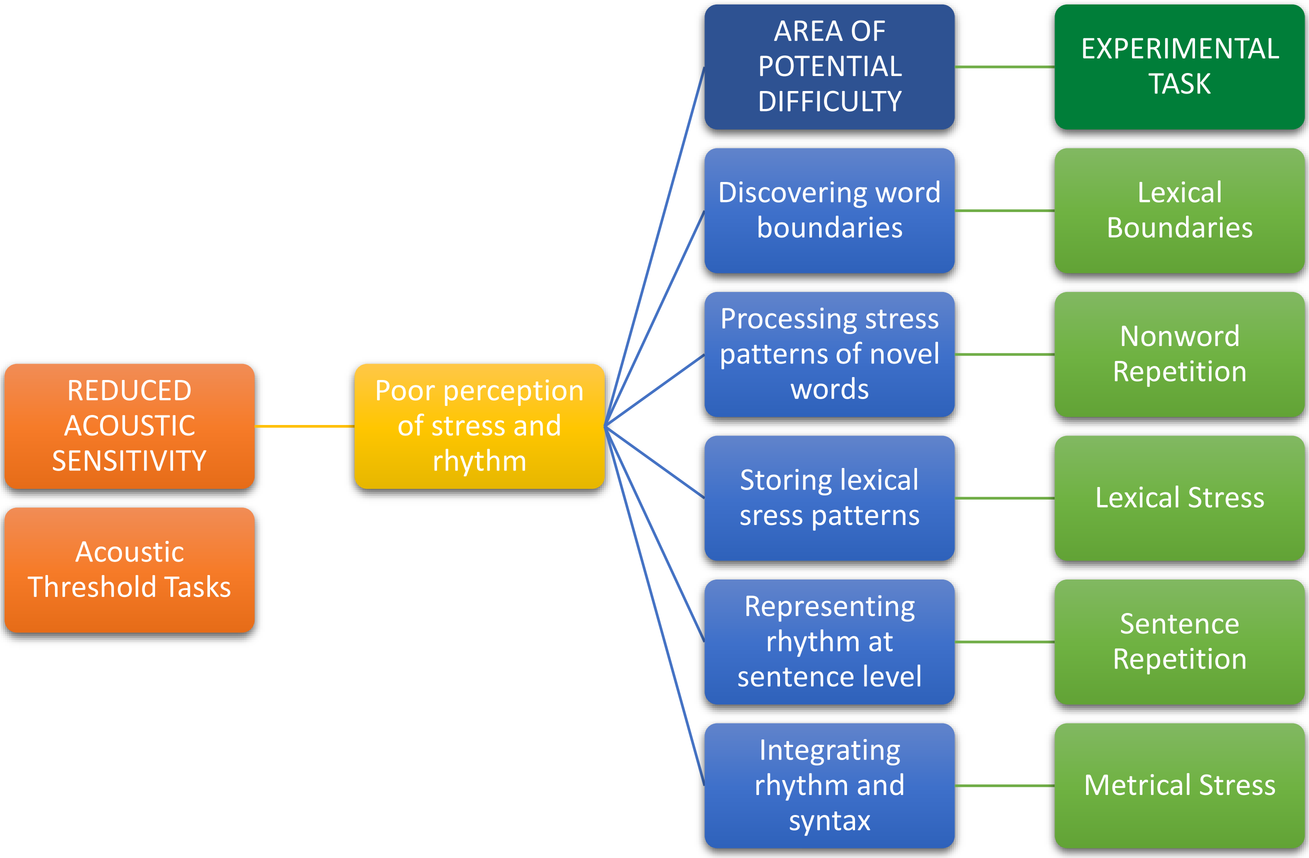


Figure 1-9 Graphic illustrating the structure of the experimental tasks

The following chapters will describe firstly, the general conception of the experiments together with participant characteristics and general methods. The thesis will then concentrate on each language area in turn, providing the theoretical basis of each individual task in detail, descriptions of the task itself, the results and a preliminary discussion of the findings.

Finally, there will be a general discussion of the task findings together with their implications for our understanding of rhythm, language and language disorder.

2 Main Study - General Introduction

We set out to investigate five areas of language processing and their relationship to language rhythm. We therefore conducted five complementary experiments, each focusing on a specific language area in which rhythmic processing is thought to play a significant role.

This section will describe the study participants and methods which were common to all the experiments, with subsequent chapters dealing with each experiment in detail.

2.1 Participants

Participants were recruited into three different groups: children with DLD, Age-Matched Children (AMC) and Younger, Language-matched, Children (YLC).

Children with DLD were recruited via schools through teacher nomination. Language status was then determined by administration of standardised language tests: (Clinical Evaluation of Language Fundamentals-3rd Edition UK standardisation (CELF-III^{UK}) (Semel, Wiig & Secord, 2000) subtests: Formulated Sentences, Recalling Sentences, Concepts & Directions; British Picture Vocabulary Scales-2nd Edition (BPVS II) (Dunn, Dunn, Whetton & Burley, 1997). Children whose score fell at or below -1.33 SD on two or more of the subtests were placed in the DLD group.

AMC children were largely recruited via the same schools as the DLD children. AMC children also completed the standardised language tests. Any AMC child whose score fell at or below -1.33 SD on any of the four subtests was excluded from the study.

Most of the YLC children attended a single school which had agreed to take part in the study for this purpose.

Participating schools covered a range of locations across Essex, Suffolk and Cambridgeshire. All were non-selective mainstream schools across state and independent sectors.

83 children were recruited in total and were divided into the three participant groups[3]: AMC (*n* = 23), DLD (*n* = 14) and YLC (*n* = 24).

---

[3] Three AMC children completed all tasks but were subsequently removed from the analysis because they had been diagnosed with dyslexia, one DLD child was also additionally diagnosed with dyslexia and so was subsequently removed. Five children were excluded due to inconclusive scores on the language tests, one child was excluded due to a hearing impairment and two children were excluded due to presence of additional developmental disorders. Four younger children with DLD (age 4:05 – 5:01) began testing, however it was discovered that their difficulties were such that they were unable to access the tasks.

2.2 Age-Matching

A one-way ANOVA revealed no significant age difference between the AMC and DLD groups (*p* = .356), whilst both AMC and DLD groups were significantly older than the YLC group (*p* = .000).

2.3 Standardised tests

2.3.1 Language tests

All DLD and AMC children completed three subtests of the CELF-III$^{UK:}$ Formulated Sentences, Recalling Sentences, Concepts & Directions. Children also completed the BPVS II, the Children's Test of Nonword Reptition (CNRep) (Gathercole & Baddeley, 1996) and the Chunking-Input prosodic subtest of the Profiling Elements of Prosody in Speech – Children's Edition (PEPS-C) (Peppé & McCann, 2003). Because of the specific input to be used in the experimental Sentence Repetition task, three subsections of the Test for Reception of Grammar (TROG) (Bishop, 1983) were also completed by all children: sections N (postmodified subject),R (relative clause), and T (Embedded sentence). As this was only a partial use of the test, no standard scores were derived, but a measure of number of items correct was obtained (maximum score of 12).

The YLC group completed CELF-III^{UK}: Recalling Sentences, BPVSII, CNRep and the TROG subsections. As they were too young for the standardisation protocol for the CELF-III^{UK}, a raw score only was obtained for language-matching purposes. The remaining CELF-III^{UK} subtests and PEPS-C were not administered so as to ease the testing burden for these younger children.

One-way ANOVAs (see Table 2-1) revealed a significant main effect of group for each of the language tests. The AMC group scored significantly more highly than the DLD group for each of the language tests, *p* = .000 (except TROG 1, *p* = .006, TROG 3, *p* = .025, PEPS-C, *p* = .035). The AMC and DLD groups therefore had significantly different language skills. This was true for both raw scores (Table 2-1) and standardised and scaled scores (Table 2-2).

Raw scores were used to investigate group differences between DLD and YLC groups. There was no significant group difference between DLD and YLC groups for the majority of the language measures, confirming that the younger, YLC group, had a similar level of language skill to the older DLD group. The only exception was the CNRep, in which the DLD children scored more poorly than the YLC group (*p* = .01).

Table 2-1 Results (Raw Scores) of standardised tests by group – one-way ANOVA and independent samples t-tests

Test	**AMC Mean (SD) [Range]**	**DLD Mean (SD) [Range]**	**YLC Mean (SD) [Range]**	***df***	***F***	***p***
Age (months)[b,c]	107.33 (16.447) [77-132]	99.71 (19.416) [66-140]	66.96 (5.78) [57-84]	2, 58	50.718	.000
Language						
Recalling Sentences[a,b]	46.38 (12.917) [26-65]	21.15 (6.950) [9-35]	24.59 (7.275) [14-45]	2, 33.536[e]	32.137	.000
Formulated Sentences[a]	31.58 (7.638) [11-42]	16.08 (5.499) [7-25]	-	1, 35	41.628	.000
Concepts & Directions[a]	25.13 (4.331) [17-30]	14.79 (6.117) [5-23]	-	1, 35	36.259	.000
BPVSII[a,b]	98.92 (21.279) [58-135]	70.43 (11.501) [56-83]	68.52 (13.601) [44-93]	2, 36.433[e]	18.535	.000
TROGN[a,b]	3.71 (.690) [1-4]	2.29 (1.383) [0-4]	3.00 (.976) [1-4]	2, 28.351 [e]	8.565	.001
TROGR[a,b]	3.63 (.576) [2-4]	2.5 (.941) [1-4]	2.77 (1.11) [0-4]	2, 28.674 [e]	11.103	.000
TROGT[a,b]	1.71 (1.459) [0-4]	0.64 (.929) [0-3]	0.77 (.869) [0-3]	2, 34.069 [e]	4.357	.025
TROG total[a,b]	9.04 (2.236) [4-12]	5.43 (2.243) [2-8]	6.55 (1.738) [4-9]	2, 57	15.702	.000
PEPS-C[a]	12.58 (2.781) [6-16]	10.64 (2.341) [6-14]	-	1, 36	4.813	.035
CNRep[a,b,d]	32.08 (3.412) [26-36]	20.38 (5.14) [15-26]	25.3 (5.321) [15-38]	2, 57	29.616	.000
Phonological Awareness						
Rhyme[a,b]	17.13 (4.963) [4-21]	7.21 (5.423) [0-17]	8.52 (4.679) [3-18]	2, 58	24.725	.000
Spoonerisms[a]	17.50 (6991) [6-29]	7.29 (5.483) [0-18]	-	1, 36	21.923	.000

Note: a) AMC > DLD b) AMC > YLC c) DLD > YLC d) YLC > DLD e) Welch's *F* and *df* used due to significant Levene's test f) as only two groups completed these tests, an additional Independent Samples t-test was carried out for these tests.

Test	**AMC** **Mean (SD)** **[Range]**	**DLD** **Mean (SD)** **[Range]**	**YLC** **Mean (SD)** **[Range]**	***df***	***F***	***p***
IQ						
Picture Completion[a,b]	18.08 (4.772) [5-24]	14.07 (4.009) [6-19]	11.18 (3.231) [4-16]	2, 57	16.529	.000
Block Design[a,b]	35.88 (13.401) [10-59]	22.64 (10.233) [8-45]	13.87 (10.872) [5-53]	2, 58	20.615	.000
Digit Span forwards[a,b]	8.46 (1.693) [6-11]	6.07 (1.207) [4-8]	7.17 (1.969) [3-11]	2, 58	8.984	.000
Digit Span backwards[a,b]	4.21 (1.25) [2-6]	3 (.961) [1-5]	2.48 (1.163) [0-6]	2, 58	13.648	.000
Digit Span total[a,b]	12.71 (1.681) [10-17]	9.07 (1.774) [6-13]	9.65 (2.656) [5-14]	2, 58	17.677	.000
Other[f]				***df***	***t***	***p***
Formulated Sentences[a]	31.58 (7.638) [11-42]	16.08 (5.499) [7-25]	-	35	6.452	.000
Concepts & Directions[a]	25.13 (4.331) [17-30]	14.79 (6.117) [5-23]	-	35	6.022	.000
PEPS-C[a]	12.58 (2.781) [6-16]	10.64 (2.341) [6-14]	-	36	2.194	.035
Spoonerisms[a]	17.50 (6.991) [6-29]	7.29 (5.483) [0-18]	-	36	4.682	.000

Note: a) AMC > DLD b) AMC > YLC c) DLD > YLC d) YLC > DLD e) Welch's *F* and *df* used due to significant Levene's test f) as only two groups completed these tests, an additional Independent Samples t-test was carried out for these tests.

Table 2-2 Results (Standard and Scaled Scores) of Standardised tests by group - independent samples t-tests and one-way ANOVAs

Test	**AMC Mean (SD) [Range]**	**DLD Mean (SD) [Range]**	**YLC Mean (SD) [Range]**	***df***	***t***	***p***
Language						
*Recalling Sentences[a]	10.38 (2.102) [7-14]	4.54 (1.941) [3-9]	-	35	8,274	.000
*Formulated Sentences[a]	10.46 (2.502) [7-15]	4.54 (1.713) [3-8]	-	35	7.597	.000
*Concepts & Directions[a]	10.74 (2.34) [8-14]	5.29 (1.49) [3-8]	-	35	7.790	.000
^BPVSII[a]	108.58 (13.292) [85-133]	89 (10.138) [64-101]		36	4.755	.000
IQ						
*Digit Span[a]	9.96 (1.574) [7-13]	6.62 (1.193) [5-9]	-	35	6.675	.000
*PC[a]	10.42 (2.518) [4-15]	8.46 (1.664) [5-11]	-	35	2.510	.017
*BD[a]	10.42 (2.888) [6-15]	8.15 (1.725) [5-11]	-	35	2.577	.014
*Mean NVIQ[a]	10.417 (2.316) [5.5-14]	8.3077 (.855) [7.5-9]	-	32.131	3.988	.000
Phonological Awareness						
^PhaB rhyme[a]	108.83 (13.53) [86-131]	82.46 (9.863) [69-102]	-	35	6.178	.000
^PhaB spoonerisms[a]	106.75 (9.755) [88-131]	91.08 (8.45) [78-111]	-	35	4.879	.000
Reading						
^BASII Word Reading[a]	104.83 (15.55) [70-131]	90 (6.76) [74-97]	-	33.998	4.06	.000
Other[d]				***df***	***F***	***p***
^BPVSII[a,b]	108.58 (13.292) [85-133]	89 (10.138) [64-101]	111.39 (11.496) [100-132]	2, 58	16.862	.000
^BASII[a,b]	104.83 (15.55) [70-131]	90 (6.76) [74-97]	115.24 (14.758) [91-145]	2, 36.878[c]	25.965	.000

Note: *) Scaled Scores *M* = 10, SD = 3; ^) Standard Scores *M* = 100, SD = 15; a) AMC > DLD; b) YLC > DLD; c) Welch's F used due to significant Levene's test d) As all three groups obtained standard scores for BPVSII and BASII an additional ANOVA was run for these tests.

2.3.2 Phonological Awareness

All participant groups completed two subtests (Rhyme and Spoonerisms) of the Phonological Assessment Battery (PhAB) (Frederickson, Frith, & Reason, 1997). The AMC group had significantly better phonological awareness scores than the DLD and YLC groups, whilst the DLD and YLC groups did not differ in their raw scores. From the standardised scores, we can see that the mean DLD standard score (82.46) on rhyme awareness was more than one standard deviation below the standardised mean (100), and that the mean group score on spoonerisms (91.08) was at the lower end of the average range, approximately one standard deviation below the mean of the AMC group. This indicates that the DLD children as a group may have some relative difficulties in phonological awareness, however these differences were not as severe as their general language difficulties.

2.3.3 IQ

All participant groups completed three subtests of the Wechsler Intelligence Scale for Children – 3rd Edition (WISC-III) (Wechsler, 1992): Picture Completion, Block Design and Digit Span. The two subtests of Picture Completion and Block Design are frequently used to give a short-form measure of non-verbal IQ (e.g. Botting & Conti-Ramsden, 2001; Fraser, Goswami, & Conti-Ramsden, 2010; Leonard et al., 2007). Digit Span was included as a measure of working memory.

The standardised scores for Picture Completion and Block Design show that the DLD group mean was within the average range, indicating that as a group they did not have depressed non-verbal IQ skills. There was nevertheless a group difference between DLD and AMC scores, with the AMC group scoring significantly more highly on each subtest. This led to a significant group difference on NVIQ (calculated from the mean of the two subtests). The groups were therefore not matched in non-verbal IQ, despite the DLD group's non-verbal IQ falling within the average range.

The YLC group were too young for the WISC-III standardisation, therefore raw scores only were used. They tended to score less well than the DLD group, however there were no significant group differences in raw scores between the DLD and YLC groups.

The AMC group had a significantly better Digit Span score than the DLD group (*p* = .000) (Raw and Scaled), whilst an ANOVA of the raw scores shows that the DLD group did not differ from the YLC group in their Digit Span scores.

2.3.4 Reading

All participant groups completed the Word Reading subtest of the British Ability Scales – 2nd Edition (BASII) (Elliott, Smith & McCulloch, 1996) as a measure of reading competence.

Due to the scoring protocols of the BASII, it is not possible to obtain comparable raw scores, therefore only standardised scores are reported. The AMC group had significantly better reading scores than the DLD group (*p* = .000), with the mean DLD score in the low average range.

2.4 Administration of Tasks

All children completed the standardised and experimental tasks individually in school. Session length varied according to child age, attention and the demands of the school day, but were typically 35 minutes in length. Most children completed the entire task battery across 5 separate sessions. Children were rewarded with stickers at the end of each session and at the end of the final session were presented with a certificate and novelty pen as a thank you for taking part.

All experimental tasks were presented on a laptop computer using Presentation software, with the children listening through Sennheiser HD650 headphones played via a UGM96 soundcard. Three tasks required a button-press response which was delivered through the laptop keyboard. Children were asked to press the button with either the pink ‘cross’ sticker or the green ‘tick’ sticker, corresponding to the ‘A’ and ‘L’ buttons of the keyboard (Figure 2-1). The Presentation software also recorded the reaction time as the time between the onset of the stimulus and the button-press. Two tasks required a spoken response, which was recorded on a PHILIPS LFH0862 digital voice recorder for later transcription.

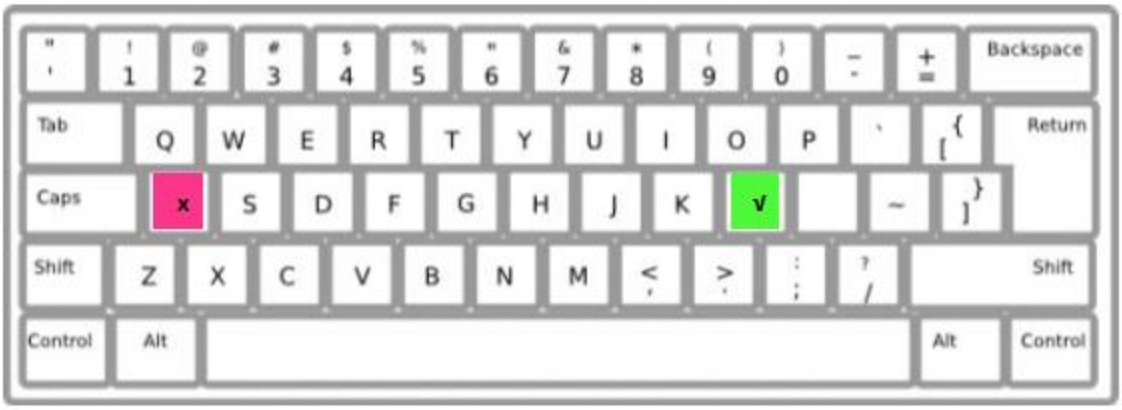


Figure 2-1 Schematic layout of keyboard

2.5 Recording of stimuli

All spoken stimuli were recorded by a female speaker of British English in a sound-proof booth, using a TASCAM DR-100 Recorder via a SHURE SM58 condenser microphone, in a manner consistent with the timing requirements of each task (see individual experiment chapters for details).The timing characteristics of all spoken tokens were verified using Audacity[4] software and adjusted as necessary.

---

[4] Audacity is an audio recording and editing program freely available from http://www.audacityteam.org/

3 Acoustic Threshold Estimation Tasks

3.1 Theoretical basis of the tasks

Acoustic analysis of stress suggests that the percept arises from a combination of four aspects of the speech signal: Rise Time, Duration, Frequency and Intensity (Greenberg, Carvey, Hitchcock, & Chang, 2003). Stressed syllables tend to have larger amplitude rise times towards the syllable nucleus, are longer, are higher in pitch and have greater intensity than their unstressed counterparts. If children with DLD have difficulty in processing any or all of these acoustic cues to stress, then it follows that their stress (and thereby rhythm) perception is likely to be impaired.

Children with DLD have been shown to have poorer sensitivity to Rise Time (see Section 1.3.1.2.2) and occasionally Frequency (see Section 1.3.1.1). In addition, children with DLD also have poorer discrimination for duration (see Section 1.3.1.3). In contrast, several studies have however found that children with DLD do not differ significantly from their age-matched peers on measures of intensity discrimination (Fraser, Goswami, & Conti-Ramsden, 2010; Mengler, Hogben, Michie, & Bishop, 2005; Richards & Goswami, 2015), although there may be difficulties for subsets of children (McArthur & Hogben, 2001).

We therefore predicted that children with DLD would have elevated thresholds for measures of rise time, duration and possibly frequency. We also predicted that elevated acoustic thresholds would correlate with poorer performance in the experimental stress and rhythm tasks.

All AMC and DLD children completed the Acoustic Threshold (AT) tasks[5].

3.2 Acoustic Threshold Task Description

The AT tasks were presented in a child-friendly format using the Dino program[6] which has been used successfully in a variety of previous studies (e.g. Corriveau, Pasquini, & Goswami, 2007; Fraser et al., 2010; Richards & Goswami, 2015).

All the AT tasks were presented in an AXB format with three tones presented per trial. The second tone (X) was always the reference tone, whilst the first (A) and third (B) were either the reference tone or a tone which differed from the reference by a stipulated amount. The status of the first and

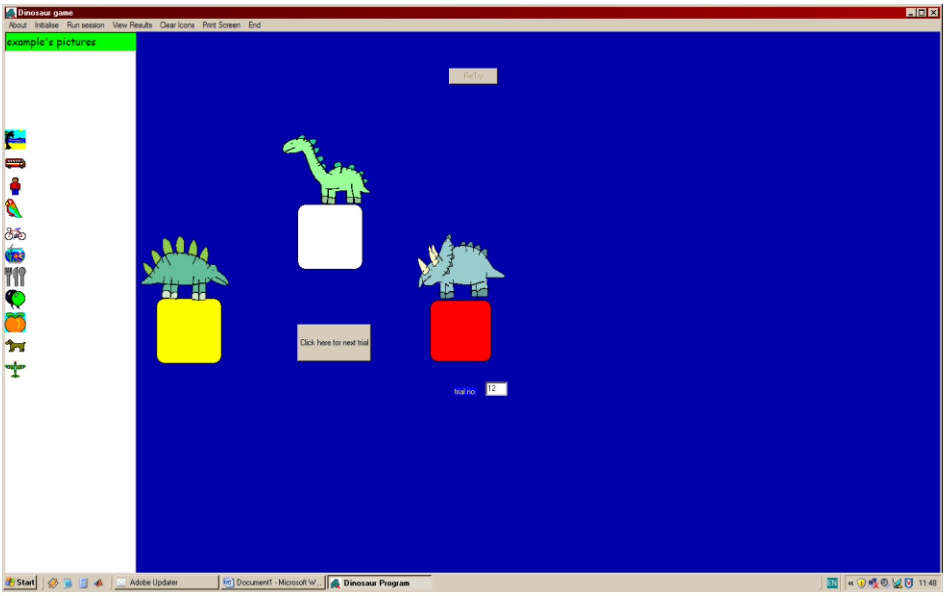
---

[5] Testing began with the YLC group, however it transpired that these 5-year-old children struggled to maintain sufficient concentration to complete these versions of the tasks and so testing for this group was discontinued. Acoustic Threshold results are therefore reported for AMC and DLD groups only.

[6] The Dino program was originally developed by Dorothy Bishop (Oxford University) and was then developed further in Cambridge by Martina Huss.

third tones (reference or target) was randomised for each trial. The AXB format was chosen in order to minimise the memory demands of the task. With three tones played, decisions can be made at two time points. If AX are different, then A is the target, confirmed when B is played. If AX are the same, then B is the target, confirmed when B is played. Previous researchers have argued that this is preferable to forced-choice format in which two tones are played with a single decision-making point (Mengler et al., 2005).

Administration of each task was the same. Children were shown an onscreen picture of three cartoon animals and told that each animal would make a noise. Their job was to listen for the animal that made the different noise. For each task, it was explained how the sound would be different (e.g. the different animal would make the longest sound). As each tone played, the relevant animal jumped up on its box. The children responded either through a mouse click on the relevant box, or by pointing. Continuous feedback was provided by the program with successful identification indicated by a positive noise and the addition of a colourful icon on the left of the screen; unsuccessful identification was indicated by a sigh. Each task was preceded by five practice trials during which the experimenter also provided live feedback and further reminders or explanations of the task as needed. The child then proceeded with the experimental trials.



trial no. 12

Figure 3-1 Example screenshot of the on-screen display during the rise time discrimination task.

The tasks were presented in a fixed order of Frequency, Intensity, RiseTime, Duration.

3.2.1 Frequency

The stimuli consisted of a series of tones each of duration 200ms played at 80.95dB. The minimum frequency was 250Hz (reference tone) and the maximum frequency 279.82 Hz. The frequency intervals between tones were of incremental intervals of 0.0513 semitones. The on-screen picture was of three elephants and the children were asked to identify which made the different, highest sound.

3.2.2 Intensity

The stimuli consisted of a series of tones each of duration 200ms and a frequency of 250Hz. The minimum intensity was 61.472dB and the maximum intensity was 80.95dB (reference tone). The intensity intervals between tones were of incremental intervals of 0.5128dB. The on-screen picture was of three cartoon mice and the children were asked to identify which one made the different, quietest sound.

3.2.3 Rise Time

The stimuli consisted of a series of tones, each of duration 800ms played at 80.95dB at a frequency of 531.25Hz. The minimum rise time was a 15ms slope (reference tone) and the maximum was a 300ms slope. The fall-off was consistent at 50ms. The minimum rise time of 15ms was always used as the reference tone and incremental intervals were set at 7.0377ms. The on-screen picture was of three cartoon dinosaurs and the children were asked to identify which one made the sound with the different, gentlest beginning.

3.2.4 Duration

The stimuli consisted of a series of tones played at 80.95dB with a frequency of 250Hz. The minimum duration was 400ms (reference tone) and the maximum duration was 595ms. The duration intervals between tones were of incremental intervals of 5.1282ms. The on-screen picture was of three cartoon sheep and the children were asked to identify which one made the different, longest sound.

3.3 Estimating Thresholds

The Dino program uses a staircasing procedure in order to estimate threshold. Trials begin with the maximum difference (i.e. between level 1 and level 40 of 40 possible levels) and initially uses a two-up, one-down procedure (i.e. two correct answers at a given difference pairing, and the program moves to a closer (harder to discriminate) difference pairing; one incorrect answer, and the program moves to a more distant (easier to discriminate) difference pairing). After the fourth reversal, this becomes a three-up, one-down procedure. Initially, pairings move by eight levels in each stepchange (e.g. move from 1:40 to 1:32), following the fourth reversal this is progressively halved to a four, two then one level difference, in order to home in on the most accurate estimation. The final threshold figure is taken as the mean level from the fourth reversal.

3.4 Results

A series of t-tests were carried out to determine any between-group differences on the Acoustic Processing measures (Table 3-1).[7]

Table 3--3-1 Results of t-tests by Group for Acoustic Threshold tasks

Task	**AMC Mean (SD)**	**DLD Mean (SD)**	***df***	***t***	***p***
Rise Time (ms)[a]	127.529 (81.81)	202.651 (65.86)	36	-2.922	.006
Frequency (semitones)[a]	0.668 (0.537)	1.391 (0.497)	34	-3.978	.000
Duration (ms)[a]	83.626 (44.065)	124.333 (51.97)	36	-2.571	.014
Intensity (dB)	-3.038 (1.516)	-3.421 (1.576)	35	.734	.468

Note: a) AMC < DLD

The DLD group had significantly higher thresholds (i.e required a greater difference to discriminate between stimuli) than the AMC group for RiseTime, Duration and Frequency. There was no group difference for Intensity.

The DLD group therefore had poorer sensitivity to three of the four acoustic cues to stress than the AMC group.

3.5 Discussion

The pattern of results obtained is similar to that found in previous studies, with the DLD group impaired in discrimination of Rise Time and Duration. This participant group also had particular difficulties with the discrimination of Frequency. Although frequency discrimination was preserved in the previous DLD sample studied by Richards & Goswami (2015), impaired frequency perception has been reported in other studies of DLD children (e.g. Cumming, Wilson, & Goswami, 2015) and so this result is not without precedent.

---

[7] Three scores were not recorded by the software – two Frequency scores (one AMC child, one DLD child), one Intensity score (one AMC child).

In comparison with their age-matched peers therefore, the children with DLD had difficulties in discriminating fine-grained changes to Rise Time, Frequency and Duration – all of which are significant acoustic cues to the detection of stress and thereby language rhythm. We therefore expected that this poorer acoustic sensitivity would relate to the experimental tasks probing aspects of language rhythm.

4 Entrainment

Before treating each experimental task individually in the following chapters, we will discuss a common factor which runs throughout the set of experimental tasks, which is the concept of entrainment.

4.1 Introduction to Entrainment

Entrainment was central to the construction of the experimental task battery, and was included across tasks as a support mechanism. We expected children with DLD to perform more poorly than the AMC controls in each of our experiments and we hypothesised that their difficulties would be related to poor processing of the acoustic cues to speech rhythm in the form of the amplitude envelope.

As a practising Speech and Language Therapist, however, I was also interested to find out whether there was a means of supporting the children in completing these tasks. Could there be a way in which rhythm could be used to provide a framework which would facilitate task completion? Dynamic Attending Theory (DAT) and neural oscillation research suggested that prior entrainment to a rhythmic structure may have a facilitating effect on language processing.

4.1.1 Theoretical Basis of Entrainment

DAT (Jones & Boltz, 1989) proposes that attentional resources make use of temporal regularities in sensory stimuli in order to predict the next occurrence of a relevant stimulus. These ‘expectancies’ focus over long time periods, whereas analytic attention focuses on lower levels of the temporal hierarchy of a stimulus. In the view of Jones & Boltz, the expectancies created by a regular temporal structure enable the construction of larger, more coherent representations; whereas when there is no clear temporal structure, focus is restricted to a succession of individual events. If we were to apply this theory to language disorder, it may suggest that children who are less able to depend on language rhythm to derive expectancies may be focusing on smaller linguistic units rather than structures across larger timeframes such as words or syntactic phrases. A future-orientated system suggests that presenting a rhythm prior to the stimulus could facilitate the directing of attentional resources to the salient point in time by deliberately inducing an attentional expectancy.

In expanding on this theory, Large & Jones (1999) suggested an ‘entrainment hypothesis’ by which internal oscillators, which they termed ‘attending rhythms’ created these expectancies, thus enabling the anticipation of future events. This insight anticipates research into neural oscillations (see Section 1.3.2) which also predicts concentration of neural and attentional resources based on rhythmic cycles set by external events. These complementary theories therefore have two central features in common: firstly, that internal rhythms are driven by the rhythmic structure of an external event; secondly that internal rhythms adjust to new rates of temporal input. This means that an external stimulus, such as the rhythmic structure of speech, could drive the creation of expectancy about the timing of the next salient event, directing attentional resources accordingly. In other words, the internal rhythm *entrains* or *synchronises* to the external rhythm. Deliberately providing an external, entraining rhythm prior to a target stimulus would therefore enable the synchronisation of internal rhythms to maximise attentional resources at the critical point.

Presentation of a stimulus within a regular rhythm has been shown to facilitate pitch judgements, with pitches presented at rhythmically expected times processed more accurately (Jones, Moynihan, MacKenzie, & Puente, 2002). Schmidt-Kassow & Kotz (2008, 2009) found that a regular, temporally predictable presentation of word onsets facilitated syntactic processing in adults, whilst presenting a rhythmic prime (in the form of a marching beat) to adults with basal ganglia lesions resulted in enhanced responses to syntactic violations (Kotz, Gunter, & Wonneberger, 2005). Przybylski et al., (2013) investigated the effect of either a regular (series of beats inducing a regular metrical pulse) or irregular musical prime on grammaticality judgements of children with DLD or dyslexia. They found that both children with DLD and TD controls had better grammaticality judgement scores after a regular prime than after an irregular prime. These results suggest that processing of auditory stimuli, including language stimuli, can be enhanced when presented within a rhythmic context.

4.1.2 Creating the Entrained Tasks

We reasoned that providing a clear rhythmic beat preceding the stimulus would facilitate accurate completion of the task. From the predictions of DAT, the preceding beat would focus attention to the salient point in time of the stimulus by creating a temporal expectancy. From a neuroscientific perspective this would arise through causing neural oscillations to reset and phase-lock to the appropriate temporal structure. This would create optimal neuronal excitability at the relevant time point and hence facilitate accurate processing. We also reasoned that a non-speech prime would contain less potentially distracting spectro-temporal information and hence provide a simplified, more easily processed, input than a speech-based prime.

We therefore created an Entrained version of each experimental task[8]. In the entrained versions, each stimulus was preceded by an entrainment track which represented either the prosodic structure of the target word (e.g. wSw) or the overall patterning of the sentence (e.g. SwSwSwSwSw). We hypothesised that performance in the entrained versions of the tasks would be enhanced compared to the standard version. Stimuli for the entrainment tracks were produced using Matlab to be commensurate with the average duration, intensity and frequency of the spoken tokens and were inserted prior to the spoken stimulus using Audacity software.

---

[8] Due to the nature of its construction, Task Five (Metrical Stress) required three iterations to complete. Having parallel Unentrained and Entrained versions would have resulted in participants completing six iterations in total, which we decided would be unduly onerous. Children therefore completed an entrained version of this task only.

5 Experimental Task 1 – Lexical Boundaries

5.1 Theoretical Basis of the Task

The first experimental task explored the processing of prosodic cues to determine word boundaries. We saw in Section 1.4.2 that the Metrical Segmentation theory suggests that the characteristic stress patterns of English may cue infants in the process of extracting word-units from continuous input. If children with DLD have reduced sensitivity to the acoustic cues which determine perception of stress, then they may have greater difficulty in distinguishing the prosodic patterns which indicate word boundaries. Such a difficulty could lead to problems in distinguishing individual lexical items and thus to slower lexical development, and indeed poor infant segmenters go on to have poorer language outcomes. It is therefore of interest to investigate whether children with DLD have difficulty in determining word boundaries from prosodic cues.

Children's prosodic skills in this regard have not received much attention, however a comprehensive task battery (PEPS-C; Peppé & McCann, 2003) was used by Wells & Peppé (2003) and Wells, Peppé, & Goulandris (2004) to explore the developmental trajectory of children's prosodic skills. The battery explores various aspects of prosody under the umbrella terms of 'chunking', 'affect', 'interaction' and 'focus' in both input and output forms. Our attention was caught by the 'chunking input' task in particular as it focused on the interaction between prosody and syntax, whereas the other three elements have pragmatic, rather than syntactic, functions. The authors defined 'chunking' as the prosodic delimitation of an utterance into units. In their task, they played children a recording of a lexical string in which the first elements were either a compound noun (e.g. chocolate-biscuits) or represented two separate lexical items (chocolate, biscuits). The children then had to indicate which of two sets of pictures matched the spoken stimulus. This task was of particular interest since it focuses on the children's abilities to use prosodic information in order to distinguish the location of word boundaries. Wells, Peppé & Goulandris argue for the difference between the stimuli sets as being one of contrasting numbers of accent groups. We would deconstruct this notion further and suggest that lower level skills are also necessary in terms of the acoustic processing of the cues which create the accent groups. An alternative way of conceptualising the difference between compound and single item nouns could be in considering the different patterning of stressed and unstressed syllables between the two types. For example, in the compound noun 'rainbow', the pattern is Sw, however for the companion single items 'rain, bow' each syllable carries primary stress. The difference is therefore cued by amplitude rise times as well as pitch contours, which combine with duration and overall intensity to create the percept of linguistic stress. The ability to distinguish between these combinations of acoustic cues would therefore underpin the creation of larger prosodic units of perception such as accent groups.

Wells & Peppé (2003) also created a parallel set of tasks in which similar stimuli to those described above were used, but this time the stimuli were laryngographs in which phonetic information was removed whilst retaining prosody. Children were played pairs of stimuli and asked to identify whether they were the same or different. When completing the laryngograph version of the chunking task, children with DLD performed significantly more poorly than typically-developing children. Children with DLD also tended to scored more poorly on the full speech version of the task although the group difference reported was marginally non-significant (p = .054).

Performance on the chunking input task correlated with measures of general language skill (TROG and CELF-R Formulated Sentences) for typically developing children (Wells et al., 2004) and for children with DLD (Formulated Sentences, CELF-R) (Wells & Peppé, 2003). For children with DLD it was the only task in the battery which did so, whilst typically developing children showed correlations between various other pragmatic prosodic functions and their language skills. This could suggest that for typically developing children, prosody is a factor which is supporting their language development across a range of competences, whereas children with DLD are making more limited use of prosody in developing their language skills.

The evidence thus far is that children with DLD have greater difficulty in using prosodic cues in distinguishing between word boundaries than do typically developing children. We hypothesised that this difficulty may result from difficulties in processing the acoustic cues of amplitude rise time and pitch which should contribute to the percept of accent groups which underlies success in this task. We further hypothesised that the presence of an entraining beat might assist children in forming a prosodic representation by providing them with the abstract template onto which the stimulus could be grafted.

5.2 Devising the Experimental Task

The inspiration for the task came from the comprehensive prosodic assessment used by Wells & Peppé (2003) and Wells et al. (2004) and was based on the 'chunking input' subtask from the PEPS-C assessment.

5.3 Creating our own version

In the PEPS-C itself, there are two types of chunking input. One is a lexically-based one in which nouns are presented either as a compound (e.g. 'cream-buns and jam') or as their individual elements (e.g. cream, buns and jam). The second type of input focused on co-ordinated adjectival phrases (e.g. ‘pink and green and white socks’, with alternative prosodic readings corresponding to meanings of i) pink socks and green and white socks; and ii) pink and green socks and white socks). We wanted to concentrate on differentiating between individual word boundaries rather than on the more complex co-ordinated adjective structure and so chose to create stimuli of the ‘two-item or three-item’ type only in our version.

5.3.1 Choosing the words

The words used for the task needed to fulfil certain semantic and morphological characteristics in order to be suitable for the paradigm. Specifically:

1. Words needed to be compound nouns in which could be split into two elements
2. Once split, each of the elements could exist as a stand-alone noun
3. The compound noun and each stand-alone noun could be readily portrayed in a picture
4. In order to exist grammatically in a list, the first element had to be an uncountable noun (i.e. would not usually be preceded by an indefinite article).
5. The second element had to be either uncountable or capable of taking a plural.
6. Both the compound noun and its stand-alone nouns should be familiar to children.

For example, ‘jellyfish’ was a suitable candidate word because it splits into two stand-alone nouns: ‘jelly’ and ‘fish’. Suitable images could then be found for ‘jellyfish’, ‘jelly’ and ‘fish’. The first element ‘jelly’ is an uncountable noun meaning it can stand in isolation in a list, whilst the second element ‘fish’ can be plural and therefore does not require an article in list format. Furthermore all three words: ‘jellyfish’, ‘jelly’ and ‘fish’ should be familiar to primary school children.

Twelve suitable compound words which fulfilled these criteria were chosen to form the task items together with two practice items.

5.3.2 Creating the stimuli

During recording of the stimuli, a beat of 750ms between stressed syllables was induced in the speaker via headphones using a metronome setting of 80bpm. There was therefore a constant inter-stressed syllable interval of 750ms. The timing of this was verified and adjusted as necessary using Audacity software. The entraining beat also had an inter-stress interval of 750ms with 250ms intervals between any intervening weak beats (corresponding to unstressed syllables). Two ‘sets’ of entraining beats were given before each stimulus. In order to separate the two sets, a ‘blank’ beat was inserted between the sets. This was the temporal equivalent of inserting an extra ‘phantom’ inter-stress interval, meaning that the second set of entraining pulses remained consistent with the 750ms pulse.

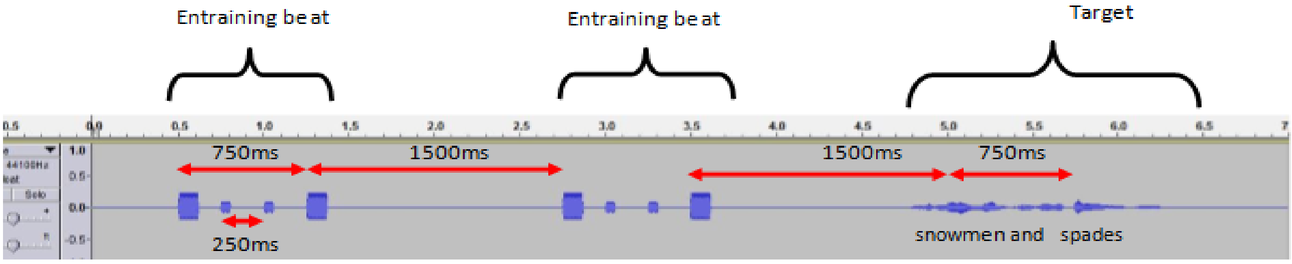


Figure 5-1 Soundwave of an example entrained two item stimulus - 'snowmen and spades'

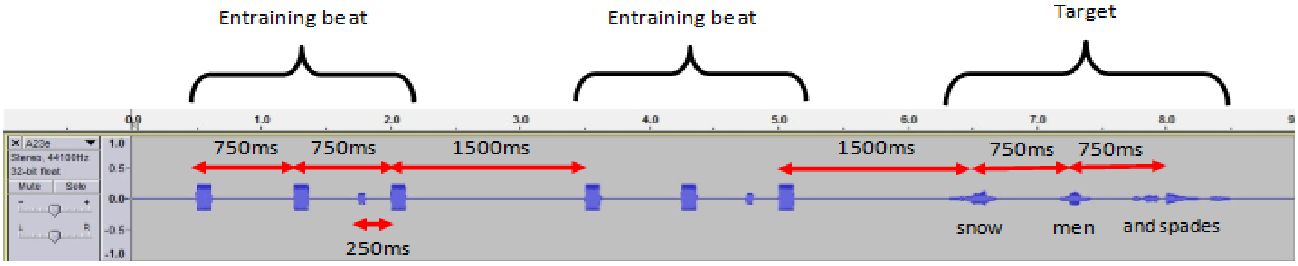


Figure 5-2 Soundwave of an example entrained three item stimulus - 'snow, men and spades'

5.4 Procedure

Two stimulus lists were created of 12 items each. Children listened to each list in either an entrained or unentrained version (for example, they might listen to List 1 Entrained and List 2 Unentrained). List and version were counterbalanced across participants as was order of presentation. Children completed each version as part of the wider task battery within a session. The complementary version was then completed in a separate session. The younger, language-matched children (YLC group) completed an Unentrained version of the task only[9][10].

Children had already completed the PEPS-C chunking input task in a previous session, so were familiar with the concept underlying the task. For each trial, a picture was displayed on the laptop, with the screen divided in two. One set of pictures was displayed on each side of the screen. The left-hand side always displayed ‘three item’ responses, the right-hand side always displayed ‘two item responses’. The ‘A’ and ‘L’ buttons on the keyboard were used as the response buttons, indicated by a plain orange sticker.

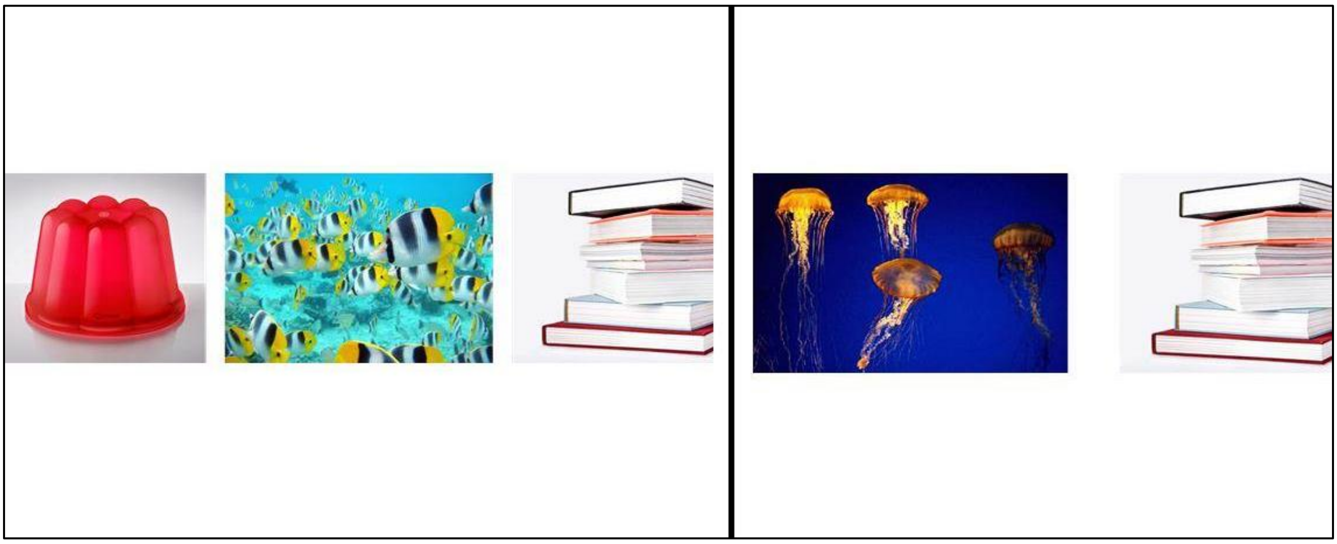


Figure 5-3 Screenshot of laptop display for stimulus pairing 'jelly, fish and books' and 'jellyfish and books'

Children were told that if they thought the words matched the left-hand picture set, they should press the left-hand button, and if they thought the words matched the right-hand picture set, they should press the right-hand button (the experimenter demonstrated by pointing). For entrained tasks, children were told they would hear some beeps first which were there to help them hear the pattern of the words. Children completed two practice items with experimenter feedback before proceeding to the experimental items.

---

[9] This counterbalancing of Entrained and Unentrained lists across participants and sessions was repeated for all the Experimental Tasks.

[10] Whilst it would have been theoretically interesting to include Entrained tasks for the YLC group as well, we were mindful of the testing burden for these younger children. In the light of results obtained for AMC and DLD groups, it was therefore decided no to gather additional Entrained data for the YLC group.

5.5 Results

5.5.1 Scores

Each child's score was calculated as the number of responses correctly identifying the stimulus. The maximum score for each list was 12.

A one-way ANOVA (DV-Score) was carried out to see if the groups differed in their accuracy levels (see Table 5-1). The AMC group was more accurate in their responses than both the DLD (p = .027) and YLC groups (p = .001), whilst there was no significant difference between the DLD and YLC groups.

Table 5--5-1 Results of one-way ANOVA for Score

Score (max 12)	**AMC Mean (SD)**	**DLD Mean (SD)**	**YLC Mean (SD)**	***df***	***F***	***p***
Score^{a,b}	11.182 (1.56)	9.143 (2.74)	8.545 (2.365)	2, 29.060	10.522^{c}	.000

Note: a) AMC > DLD b) AMC > YLC , c) Welch's *F* used due to significant Levene's test

Our prediction that the DLD group would be less accurate than typically developing children in detecting the word boundaries demarcating lists into groups of two or three items was therefore supported by the results.

5.5.2 Directions of bias

5.5.2.1 d'

It was interesting to know whether there was a systematic tendency among the groups to prefer a particular response. To that end a measure of d' was taken using 3-item-Target, 3-item Response as the hit rate, and 2-item-Target, 3-item-Response as the false alarm rate.

Table 5-2 Results of one-way ANOVA for d'

d'	**AMC Mean (SD)**	**DLD Mean (SD)**	**YLC Mean (SD)**	***df***	***F***	***p***
d'^{a,b}	2.396 (.716)	1.452 (1.262)	1.184 (1.09)	2, 29.023	10.54^{c}	.000

Note: a) AMC > DLD b) AMC > YLC , c) Welch's *F* used due to significant Levene's test

The AMC group had a significantly higher d' score than the DLD group (p = .026) and the YLC group (p = .001) whilst the DLD and YLC groups did not significantly differ on their d' score. This indicates that the responses of the DLD and YLC groups were more consistently biased than those of the AMC group.

5.5.2.2 *Preference for two or three item responses*

In order to discover in which direction this bias lay (towards ‘two items’ or towards ‘three items’) a calculation was made of the number of times each child pressed the ‘two item’ and ‘three item’ response buttons (no bias would result in a score of 6 for each).

Table 5-3 Mean number of responses of 'two items' and 'three items'

Response	**AMC** **Mean (SD)**	**DLD** **Mean (SD)**	**YLC** **Mean (SD)**
‘two items’	6.455 (1.654)	7.857 (2.35)	8.182 (2.039)
‘three items’	5.545 (1.654)	4.143 (2.35)	3.818 (2.039)

A 3 x 2 repeated-measures ANOVA (Group [AMC, DLD, YLC] x Number of Items [two, three]) revealed a significant overall effect of item number (*F*(1, 55) = 31.606, *p* =.000, with children being significantly more likely to choose the ‘two item’ response. There was also a significant group*item number interaction, *F*(2) = 4.562, *p* = .015, indicating that the pattern of responses varied significantly between the groups.

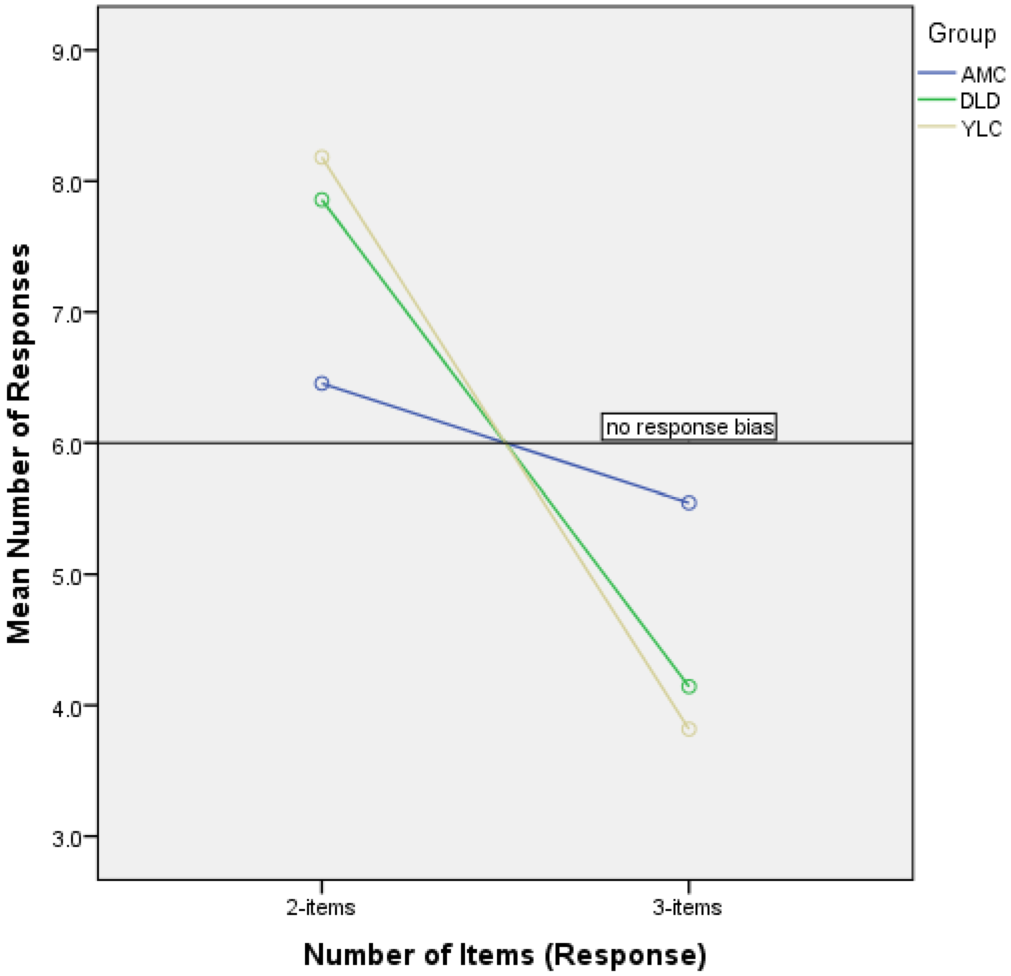


Figure 5-4 Graph showing the mean number of responses by group for two- and three-item responses. The solid black line indicates where a 'no response bias' line would fall.

Figure 5-4 shows the mean number of responses for each item number and illustrates the nature of the interaction. All groups have a tendency to prefer 'two items' as a response, however for the AMC group this is small and non-significant. On the other hand, the bias for the DLD and YLC groups towards 'two items' is clear.

The repeated-measures ANOVA was then conducted for each group individually. There was no significant difference between the number of 2-item and 3-item responses for the AMC group, *F*(1, 21) = 1.661, *p* = .211, however there was a significant difference for the DLD group, *F*(1, 13) = 8.753, *p* = .011, and the YLC group, *F*(1, 21) = 25.2, *p* = .000, with significantly more '2-item' responses than '3-item' responses.

Children from the AMC group were therefore able to accurately distinguish stimuli containing two or three items. Children from the DLD and YLC groups were less accurate and their errors displayed a preference for responding that stimuli were composed of two items.

5.5.3 Reaction Times

Reaction Times (RTs) were recorded by the Presentation software as the time between the onset of the stimulus playback and the response button being pressed in units of a tenth of a millisecond (0.0001s).

A further adaptation was made in order to avoid individual, unrepresentative scores affecting the final results. For each child, the mean and standard deviation of their RTs were calculated. Any individual time which lay outside 3 standard deviations of the mean was discarded and the child's mean time re-calculated[11]. In practice, only for one child did this make a difference to their final mean RT.

5.5.3.1 Effect of Group on Reaction Times

A one-way ANOVA (DV – RT) was conducted in order to see if the groups differed in their RTs. There were no significant group differences in the RTs, despite a general tendency for the DLD and YLC responses to be slower than the AMC group.

Table 5-4 Results of one-way ANOVA for Reaction Times

	AMC Mean (SD)	**DLD Mean (SD)**	**YLC Mean (SD)**	***df***	***F***	***p***
Reaction Time (.0001s)	40268 (8988)	47540 (18300)	52095 (24765)	2	2.264	.114

Speed of reaction therefore did not vary between the participant groups.

5.6 Effect of Entrainment

We predicted that listening to an entraining beat would have a facilitating effect on task performance, resulting in higher scores in the Entrained version of the task.

---

[11] This procedure was carried out for each experiment in which Reaction Time was recorded.

5.6.1 Effect of Entrainment on Score

To investigate the effect of Entrainment on Score, we used a repeated-measures ANOVA (2 x 2 – Group [AMC, DLD[12]] x Entrainment [Unentrained, Entrained]; DV-Score). There was no significant effect of Entrainment on score achieved, *F*(1, 33) = .555, *p* = .462.

Table 5-5 Unentrained and Entrained Scores for AMC and DLD groups

Score (max 12)	**AMC**	**DLD**
	Mean (SD)	**Mean (SD)**
Unentrained Score	11.182 (1.56)	9.143 (2.74)
Entrained Score	10.833 (1.76)	9.308 (2.14)

We had hypothesised that the presence of an entraining beat would lead to higher scores in the Entrained condition, however there was no significant difference in accuracy between the two conditions. Entrainment did not therefore result in better accuracy in determining word boundaries.

5.6.1.1 Effect of Entrainment on Reaction Times

We also predicted that listening to an entraining beat might result in faster RTs in the Entrained version of the task.

In order to compare between Entrained and Unentrained versions, the timeframe of the entraining portion was deducted from the recorded times for the entrained task – see Figure 5-5.

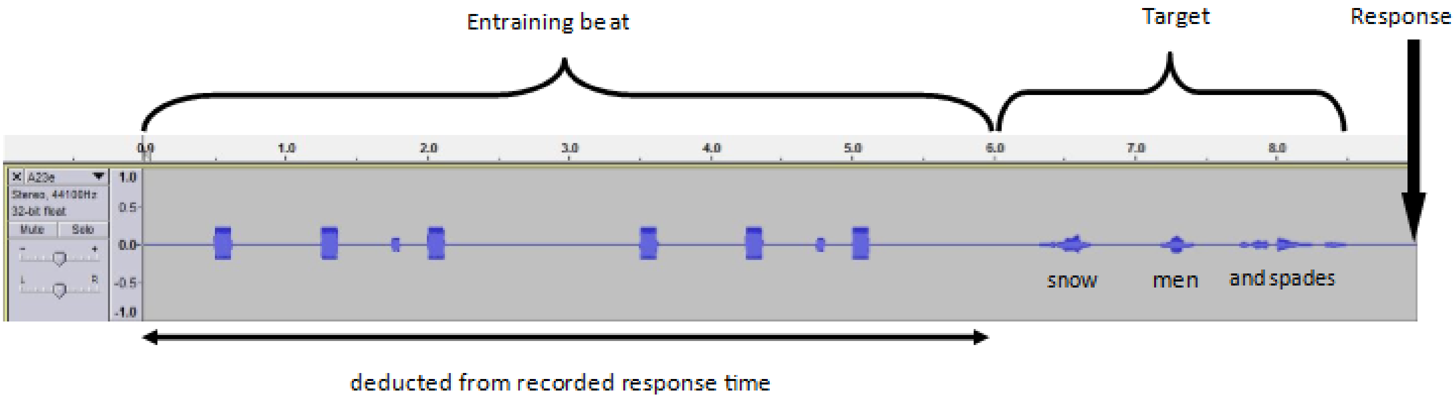


Figure 5-5 Soundwave illustrating timeframe of entraining segment deducted from response time in Entrained condition

---

[12] Recall that YLC children did not complete Entrained versions of tasks

Table 5-6 Reaction Times for Unentrained and Entrained versions of the task

Reaction Time (.0001s)	**AMC Mean (SD)**	**DLD Mean (SD)**
Unentrained	40268 (8988)	47540 (18300)
Entrained	40123 (13197)	39218 (13001)

A repeated-measures ANOVA (2 x 2 - Group [AMC, DLD] x Entrainment [Unentrained, Entrained]; DV- RT) found a significant main effect of Entrainment, *F*(1, 33) = 4.986, *p* = .032, with Entrained reaction times being faster than Unentrained times. The Group*Entrainment interaction was not significant (*F*(1) = 2.381, *p* = .132, however visual inspection of the graph (Figure 5-5) indicates a trend towards the effect of Entrainment on RT being greater for the DLD than the AMC group.

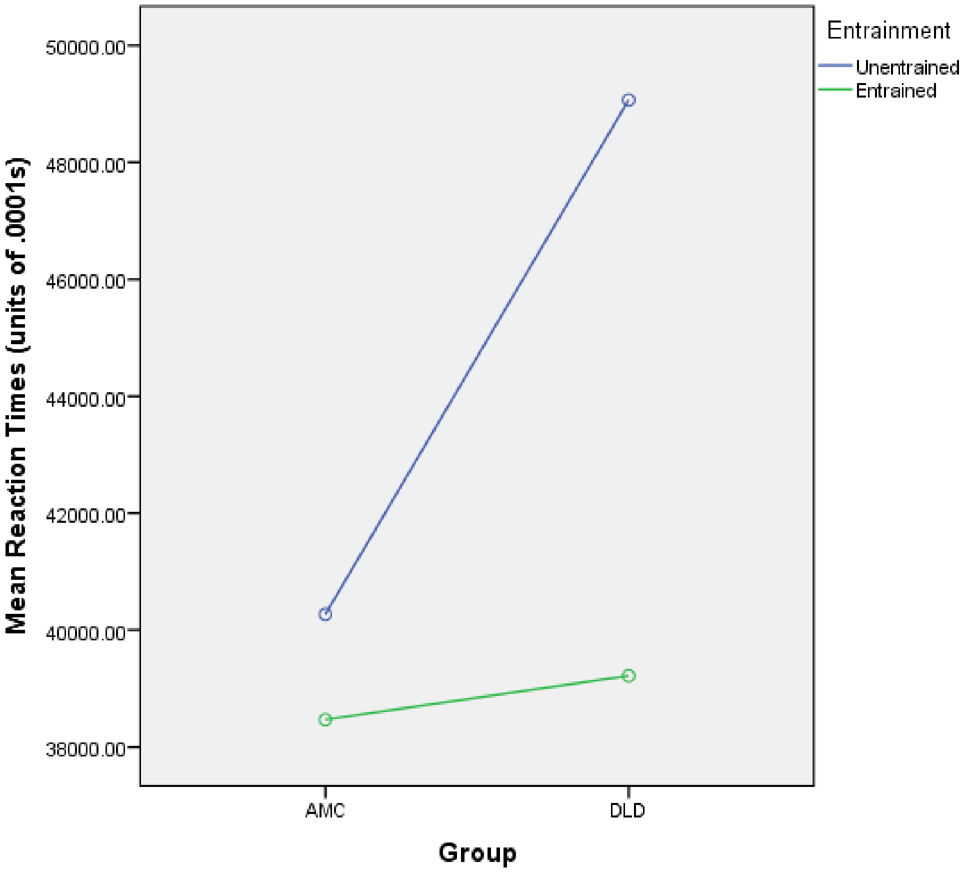


Figure 5-6 Graph showing mean reaction times by group for unentrained and entrained versions

Further repeated-measures ANOVAs for each group individually revealed no significant effect of Entrainment on RT for the AMC group, *F*(1) = .419, *p* = .524, whilst the effect of Entrainment for the DLD group approached significance, *F*(1) = 4.011, *p* = .068.

Listening to an entraining beat therefore resulted in faster reaction times overall, however this was not replicated when examined at the individual group level. The effect of entrainment seems to have been greater for the DLD group, with Figure 5-6 illustrating that their response times quickened to approach those of the AMC group in the Entrained condition. However, this must be treated with caution as the interaction did not reach significance.

5.6.1.2 Summary of Effect of Entrainment

We had predicted that listening to an entraining beat would result in more accurate task performance with greater speed of response times, however our data did not show any difference in accuracy between the Entrained and Unentrained versions of the task. We did find an overall effect of Entrainment on reaction times, with Entrained responses being quicker than those in the Unentrained version. This overall response appeared to be driven by a stronger, but non-significant, effect of Entrainment for the DLD group. Listening to an entraining beat therefore offers some support for speed, but not accuracy, of response in determining lexical boundaries.

5.6.2 Relationship with Acoustic Processing tasks

We predicted that any difficulties the DLD group exhibited in this task would be related to their acoustic processing skills. We therefore conducted a correlation analysis to investigate whether there was a statistical relationship between performance on the AT tasks (see Chapter 4) and performance on the experimental task.

Table 5-7 Pearson correlations (r) one-tailed for AMC and DLD groups for Acoustic Processing Thresholds and Lexical Boundaries Scores

Task	Duration	Frequency	Intensity	Score
Rise Time	.340*	.795***	-.227	-.437**
Duration	-	.406**	.016	-.162
Frequency	-	-	-.206	-.384*
Intensity	-	-	-	.363*

Note: * *p* < .05; ** *p* < .01; ****p* < .001

There was a significant correlation between rise time threshold and score on the experimental task (*p*= .004). The inverse relationship indicates that the smaller the threshold value, the better the score on the task i.e. that children who were able to distinguish between smaller differences in rise time tended to score better in Lexical Boundaries. There was a similar significant inverse relationship between Frequency threshold and task score (*p* = .012). There was a significant positive relationship with Intensity (*p* = .016). Because Intensity was calculated as a negative figure, this also indicated that smaller thresholds correlated with better task performance[13].

Smaller thresholds in Rise Time, Frequency and Intensity were therefore related to better task performance.

Further to this, a regression was conducted to explore the unique variance in performance accounted for by the predictors of Age[14], NVIQ and each AT measure in turn. Step 1 was always entered as Age (months), Step 2 as NVIQ and Step 3 as the AT measure[15]. Overall task score was entered as the dependent variable.

---

[13] Scatterplots illustrating the relationship between AT tasks and all experimental tasks are provided in Appendix B.
[14] Scatterplots illustrating the relationship between Age and Task Score are provided in Appendix E.
[15] Scatterplots illustrating the relationship between Age and AT Threshold are provided in Appendix D.

Table 5-8 Results of Regressions exploring the unique variance in overall score accounted for by Age (months), NVIQ and AT measures (AMC and DLD groups)

	b	SEb	β	ΔR^2	*p*
Model 1					
Age	.070	.019	.535	.386	.000
NVIQ	.013	.150	.013	.006	.571
Rise Time	-.007	.004	-.254	.053	.094
Model 2					
Age	.068	.021	.535	.354	.000
NVIQ	.079	.178	.080	.009	.515
Duration	-.001	.006	-.033	.001	.828
Model 3					
Age	.077	.020	.591	.386	.000
NVIQ	.078	.155	.075	.006	.571
Frequency	-.209	.584	-.064	.003	.722
Model 4					
Age	.072	.019	.554	.382	.000
NVIQ	.087	.148	.085	.006	.594
Intensity	.304	.203	.210	.042	.146

Age was a significant predictor of performance in this task, accounting for between 35.4 and 38.6% of unique variance (ΔR^2 range = .354 -.386, *p* = .000) in the models. In contrast, NVIQ did not account for significant changes in variance (*p* range = .515 - .571). The AT tasks contributing the largest unique variance were Rise Time (5.3%) and Intensity (4.2%). This matches the correlational findings, however neither of these contributions were significant once age and NVIQ were controlled for (*p* = .094, .146 respectively).

5.7 Summary

The Lexical Boundaries task was designed to investigate whether children with DLD were able to judge lexical boundaries using only prosodic cues. We found that the DLD group were significantly less accurate than the AMC group in judging the number of items presented in the list, with an error-bias towards responding that the stimuli contained 'two items'. In this respect, the response pattern of the DLD group resembled that of the younger, YLC, group. These results therefore lend support to the idea that children with DLD may have more difficulty than their peers in responding to the prosodic cues indicating word boundaries, and that TD children become better at using these cues as language development progresses.

We had also predicted that listening to an entraining beat would produce a beneficial scaffolding effect for the ensuing prosodic structures, thus facilitating accurate judgements between two-and three-item stimuli. We found that whilst the entraining beat did not significantly affect accuracy, there was a beneficial effect on speed of response with entrained responses faster overall than those in the unentrained condition. It therefore seems that entrainment may facilitate speed of response rather than overall accuracy in making word boundary judgements. This facilitative effect tended towards being greater for the DLD group, although not significantly so.

Furthermore, we predicted that success in the experimental task would be related to measures of acoustic thresholds. We found a relationship between sensitivity to Rise Time, Frequency and Intensity and performance on the task, however the regressions demonstrated that none of these relationships retained significance after controlling for age.

5.8 Discussion

We had predicted that children with DLD would be less accurate at detecting word boundaries indicating divisions into two- and three- item lists than the AMC group. The analysis of the raw scores confirmed this prediction as the DLD group were significantly less accurate in their judgements than the AMC group in both the Unentrained and Entrained conditions. This is similar to Wells & Peppé's (2003) result in which children with DLD were significantly poorer in the laryngograph version of the chunking task, and where the group difference in the full speech version was marginal (*p* = .054). Children with DLD therefore seem to be consistently poorer at judging word boundaries on the basis of stress cues than TD children.

Consideration of the acoustic cues differentiating the two stimulus conditions may give us some explanation of why this group difference between AMC and DLD children occurs. Essentially, the judgement in this task depends on deciding whether the second element forms part of the first word, or whether it is a separate word in its own right. Close inspection of the soundwaves of two- and three- item lists suggest that two of the major acoustic cues for making this judgement are in the characteristics of the peaks of the amplitude envelope and in the temporal distribution of those amplitude peaks.

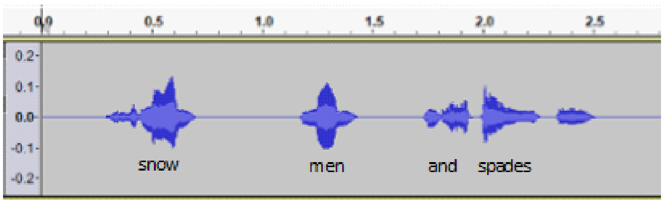


Figure 5-7 Soundwave of three item 'snow, men and spades'

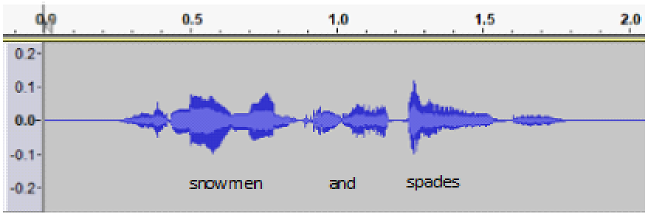


Figure 5-8 Soundwave of two item 'snowmen and spades'

In terms of temporal spacing, there is a clear difference in pattern between the two alternative readings. In Figure 5-7, there is 750ms between the first and second syllables (‘snow’ and ‘men’) including an extensive period of silence demarcating their separation. In contrast, in Figure 5-8, the two syllables occur contiguously without a clear demarcation break. It is possible that the ability to accurately judge duration of auditory events may affect the capacity to draw on temporal demarcations in order to make judgements. If so, we would have expected that the DLD children, who had higher duration thresholds in the acoustic processing tasks, would find this task harder. Neither the correlations nor the regressions, however, showed a relationship between durational processing and success in this task, suggesting that duration may not be the primary cue that is being used.

In contrast with duration, level of rise time threshold was highly correlated with task performance for the AMC group. The relevance of rise time as a cue can be seen from examining the height and slope of the amplitude peaks. The height of the amplitude peak denoting the syllable ‘men’ is greater in Figure 5-7, with a more sharply defined rise time. This is a cue that this syllable is stressed, and therefore the beginning of a new word. In comparison, the syllable ‘men’ in Figure 5-8 has a lesser peak with a more gradual onset, providing a cue that this syllable is unstressed and therefore part of the larger word ‘snowmen’. Children with DLD had significantly higher thresholds in the rise time threshold estimation task, meaning that they found it harder to make distinctions between differing amplitude rise times. Higher rise time thresholds could be detrimental to success in this task as the two differing peaks may be perceived as equivalent, therefore eliminating the usefulness of the amplitude cue. Lack of sensitivity to small changes in rise time may therefore have contributed to the DLD group’s relative lack of success in the experimental task, with the regressions showing that Rise Time thresholds accounted for a unique 5.3% of variance in task score after controlling for the significant variance accounted for by age.

The third acoustic cue to word boundaries is pitch contour. In Figure 5-9 we can see the different range and directions of the pitch trajectories (blue line) in the ‘snowmen’ and the ‘snow, men’ versions.

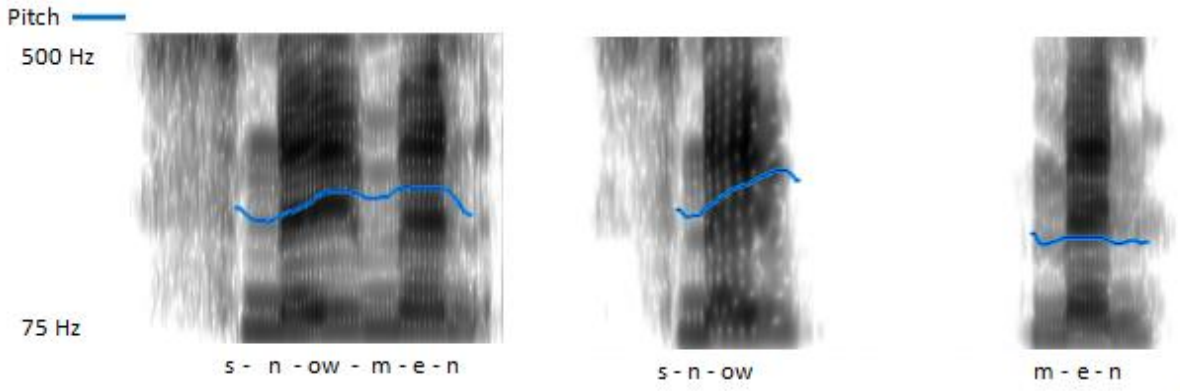


Figure 5-9 Spectograms of 'snowmen' and ‘snow, men’ superimposed with pitch (blue line)

The DLD group had elevated thresholds for detecting frequency change compared with the AMC group and so they may also have had difficulty in detecting the different pitch contours cueing the word boundary information, although Frequency threshold did not account for significant additional variance in the regressions.

There was also a significant correlation between Intensity thresholds and task performance, indicating that this cue may also have affected task completion, accounting for 4.2% of variance in addition to that accounted for by age and NVIQ.

The greater levels of differentiation in the signal required by the DLD group before they are reliably able to distinguish between acoustic properties may therefore be leading to difficulties in reliably interpreting aspects of linguistic prosody which rely on those acoustic properties for creating differing percepts.

The direction of this difficulty was hinted at by the discovery that, when making errors, children in the DLD and YLC groups were most likely to respond that the ‘three item’ stimuli contained two items, suggesting that they were having more difficulty in interpreting the rise time and pitch cues separating the first syllables into separate words.

We had predicted that providing an entraining rhythm might assist children in correctly judging the lexical composition in the stimulus by providing an acoustic cue on which they could scaffold the linguistic input. This did not prove to be the case, however, as neither AMC nor DLD groups showed a difference in accuracy between the two conditions. Nevertheless, some of the AMC participants verbalised that they were making active use of the entraining period during the experiment.

Participant 2 (AMC) said ‘I’ve worked out a pattern with the beeps’, indicating that he was fully aware of the supportive role that the beeps were playing; Participant 32 (AMC) said ‘I can tell by the beat’ and had his finger ready with his decision by the end of the entrainment section i.e. he had already worked out the answer from the entraining rhythm alone. Other participants were observed to be nodding or tapping their fingers to the beat, suggesting that they were actively engaged with the entrainment segment of the task. Despite inter-group differences, overall success in this task was relatively high, perhaps rendering the need for extra support redundant. It could be that an accuracy difference would be found with a task that presented a greater basic level of challenge.

Alternatively, the answer could lie in the nature of the entraining rhythm. In order for the entraining rhythm as a whole to assist in accuracy, children would have to extract and retain the full template (e.g. Sw w S for SNOWmen and SPADES), or be actively counting Strong beats in order for it to provide a helpful scaffold. It could be that provision of the whole rhythm to be extracted (as opposed to a consistent temporal cue) is therefore not as helpful as expected. Entrainment did, however, demonstrate an effect on reaction times. Responses in the entrained condition were significantly faster than in the unentrained condition and this was particularly true for the DLD group. This suggests that locking on to a clear entraining beat, although not affecting accuracy per se, may result in more rapid and efficient processing of speech stimuli.

In terms of theories of language processing, it seems likely that a child who is able to entrain effectively to the prevailing rhythmic pattern of speech will be able to process that speech more efficiently than those who cannot. Since speech creates a unidirectional, continuously varying signal, speed and efficiency are likely to be key to effective processing. Although the reaction times between AMC and DLD groups were not significantly different overall, the times from the DLD group tended to be slower. Under entrained conditions, however, the mean DLD reaction time was similar to that of the AMC group in the unentrained condition. This suggests that the entraining beat may have provided a compensatory mechanism for the DLD group.

In this experiment, the children with DLD were less accurate in their judgements of whether rhythmic stimuli were composed of two item or three item lists than were AMC children. The provision of an entrainment beat did not significantly affect the accuracy of responses, however it resulted in faster responses, suggesting a role for entrainment in facilitating efficient processing of speech.

6 Experimental Task 2 – Nonword Repetition

In Section 1.4.3, we saw that infants use rhythmic information in order to discriminate between words and to develop their knowledge of word-like structures. Over the first year, segmental information is grafted on to initial rhythmic groupings to create word-like percepts (Morgan & Saffran, 1995). These insights indicate that rhythm is at the core of how infants process novel words to which they are exposed. If rhythm does indeed provide the skeleton for representing phonological detail, then a difficulty in processing the rhythm of novel words is likely to lead to poorer phonological representations of those words

A task which has often been used as a measure of how successfully older children are able to process novel words is Nonword Repetition (NWR). The reasoning behind this is that NWR is thought to mimic the task faced by any language-learner in coming across a word for the first time. In essence, any new word is a 'nonword' – a string of unrelated phonemes and syllables – until the nonword acquires additional semantic and syntactic specifications.

Investigating the role of rhythm in NWR may therefore lead to insights into how the rhythmic structure of target words impacts on repetition skills in DLD, and thus into how rhythm may be influencing the processing of novel real words and vocabulary for these children.

6.1 Theoretical basis of the task

This section will first provide an outline of previous research in the field of NWR, before considering the specific role that rhythm might play in the NWR process.

6.1.1 Nonword Repetition, Language and Language Disorder

NWR has been extensively studied by Gathercole and colleagues, from whose initial studies arose the hypothesis that NWR tasks were a direct measure of phonological short-term memory (PSTM), that is the hypothesised component of working memory dealing with the on-line retention of verbal information (e.g. Gathercole & Baddeley, 1989, 1990).

Performance in NWR is therefore thought to indicate skills critical for the long-term acquisition of new phonological items (Gathercole, 1995). There is evidence for a strong relationship between NWR skills and vocabulary learning in children (Gathercole & Baddeley, 1989, 1990; Gathercole, Willis, Emslie, & Baddeley, 1991; Roy & Chiat, 2004). NWR performance also correlates with wider language skills such as spoken language measures in pre-schoolers (Adams & Gathercole, 1995, 2000; Edwards & Lahey, 1998) and with components of other standardised language assessments such as the CELF (Bishop, North, & Donlan, 1996; Conti-Ramsden, 2003)

The relationship between NWR and language learning has led to a burgeoning interest in NWR performance in children with DLD. Gathercole & Baddeley (1989) investigated NWR skills in children with DLD and found they had significant difficulties with the task, with scores below both age-matched and language-matched controls. Since then, numerous studies have replicated the finding of NWR difficulties in DLD (e.g. Bishop et al., 1996; Chiat & Roy, 2007; Dollaghan & Campbell, 1998; Edwards & Lahey, 1998; Marton & Schwartz, 2003; Munson, Kurtz, & Windsor, 2005).

Difficulties with NWR cannot always be attributed purely to a weak PSTM, however. Archibald & Gathercole (2007) found that whilst the DLD children in their study had poor Digit Span scores, these scores could not account for the disproportionate difficulties the children had with NWR, despite high levels of correlation between the tasks. In their 2006 study, Archibald & Gathercole found that children with DLD had difficulties with both short-term and verbal working memory tasks leading them to conjecture that children with DLD experience ‘double memory jeopardy’, as their short-term memory difficulties lead to problems in learning the phonological structure of language, whilst working memory difficulties result in wider processing problems in dealing with verbal material. They concluded that it was this combination of difficulties which resulted in a wide-ranging language impairment.

The consistent finding of NWR difficulties in children with DLD has led researchers to propose this task as a reliable clinical marker of language difficulties. Dollaghan & Campbell (1998) demonstrated that NWR performance successfully separated children with DLD from TD children, whilst Ellis Weismer, Tomblin, Zhang, Buckwalter et al. (2000) concluded that NWR (in conjunction with other measures) could identify children as having DLD. Conti-Ramsden (2003) examined the specificity and sensitivity of four different language tasks and found NWR to be the best marker of DLD in 5-year-old children. Difficulties with NWR have also been found even in children whose language difficulties have largely resolved (Bishop et al., 1996; Conti-Ramsden, Botting, & Faragher, 2001). Twin studies have also shown that NWR ability is highly heritable, thereby suggesting that NWR performance could be a reliable behavioural marker for heritable language impairment (Bishop et al., 1996).

Children with DLD therefore have particular difficulties with carrying out NWR tasks which go beyond simple PSTM (i.e. Digit Span). Difficulties with NWR also correlate with more general language difficulties, suggesting that there is more to the task than pure PSTM as Gathercole &

Baddeley had initially suggested. There has therefore been a gradual realisation that in terms of task construction, not all nonwords are equal. NWR performance can be affected by aspects such as phonotactics (are the phoneme sequences allowable in the language), (Munson et al., 2005) and ‘wordlikeness’ (the extent to which the nonwords sound like they could be a real word) (Gathercole, 1995). The explanation for these effects is considered to be that children do not arrive at the repetition task as a blank canvas, but instead bring their accumulated linguistic knowledge to bear on the task. Prior knowledge of language rules (phonotactics) or wider vocabulary (wordlikeness) has a positive influence on task performance. In the light of the evidence showing that NWR draws on accumulated language rules, Baddeley (2003) included interfaces with long-term memory and language systems within his adapted working memory model. Since prior language knowledge can influence NWR, we considered it possible that a further influence could be familiarity with the rhythmic structure of words.

6.1.2 NWR and Rhythm

The idea of a role for rhythm as part of the language competence drawn on in NWR is not, itself, new. Back in 1991, Snowling, Chiat, & Hulme stated that “nonword repetition is a complex psycholinguistic task that undoubtedly engages a child’s existing knowledge of the phonological, *including prosodic*, structure of language” (p371 – my italics). Nonetheless, few studies since then have considered the role of rhythmic structure as part of NWR investigations.

Steps to rectify this were taken by the researchers developing the Early Repetition Battery (Seeff-Gabriel, Roy, & Chiat, 2008). They investigated the repetition of 1-3 syllable words by typically developing preschool children, whilst systematically varying the target prosodic structure and found that prosodic structure impacted on repetition patterns. Syllables carrying primary stress were well preserved, whilst post-stress unstressed syllables occurring within a trochaic foot (i.e. Sw) were better preserved than unstressed syllables occurring in pre-stress positions. Syllables to which the authors assigned secondary stress (e.g. ‘saur’ in ‘dinosaur’) appear to have fallen between the stressed and post-stress values of accuracy, however they did not specify whether the differences were significant (Roy & Chiat, 2004). Both stress status and the location of unstressed syllables relative to stressed syllables therefore appear to be influential in determining repetition accuracy for preschool children. This suggests that the overall stress *pattern* of a word as well as the stressed and unstressed status of individual syllables may affect repetition responses.

There has been equally little investigation of how rhythmic stress structure may affect NWR in children with DLD, however Chiat & Roy (2007) investigated NWR in clinically-referred 2- and 3-year-olds, using the same stimulus materials as used previously with TD children (Roy & Chiat, 2004). They found that the rate of syllable loss (their means of measuring error) was considerably higher in the DLD group than in TD children. Whilst the overall impact of rhythmic structure was similar between the groups, one striking feature of the DLD children's responses was the omission of syllables in positions which were almost never lost in typical children. These were stressed syllables and post-stress syllables in two-syllable words. This suggests that some of the rhythmic features which may serve to 'anchor' repetition in TD young children may be weaker and more vulnerable in children who are experiencing difficulties with language.

A further study with older children is suggestive regarding the possible role of rhythm in repetition tasks. Archibald & Gathercole (2007) asked their TD and DLD participants (age range 7;6 – 13;0) to perform both a NWR task and a serial recall task in which the items to be recalled were the same syllables as had occurred in the nonwords, but in list, rather than word, form. They found that whilst the DLD group performed worse than controls on both tasks, the effect size for this was much greater in the nonword condition than in the serial condition. In other words, the DLD children exhibited a greater deficit for repeating nonwords than for repeating serial syllables with the same phonemic content. The authors attributed this disparity to 'a further ability that is specific to the repetition of novel multisyllabic phonological forms' (p.923). One such property unique to multisyllabic words is rhythmic structure. Adults also repeat nonwords more accurately when they are presented with a prosodic structure (Archibald, Gathercole, & Joanisse, 2009) supporting a role for prosody as a facilitator of accurate recall.

Since NWR is influenced by prior language knowledge, we hypothesised a role for rhythm which would explain the Archibald & Gathercole (2007) results. It is possible that the TD children were able to impose a rhythmic structure on the serial syllables by deriving this from their long-term memory store and combining the stored template with the novel incoming stream to scaffold their retention of the nonwords. This would be consistent with later models of working memory (e.g Baddeley, 2003) in which long-term knowledge is combined with on-line input in a redintegration process within the episodic buffer to create a new 'episode'. If children with DLD have less developed stored templates of potential rhythmic patterns, then they will be less able to call on these to support new episode creation.

Consideration of prior rhythmic templates could also explain a discrepancy noted in Chiat & Roy's 2007 work with preschool children with DLD. They had previously found that TD children repeated real words more accurately than nonwords, but this occurred for a much smaller percentage of the clinical sample. One explanation could be that TD children were able to call on robust rhythmic templates to support repetition of real words, whereas DLD children could not. Furthermore, the

DLD group had a higher rate of syllable loss for real words than for nonwords. This may suggest that for those words which the DLD group recognised, they were calling on incomplete representations, thereby producing a pre-stored incomplete repetition from their store, whereas for nonwords they were more reliant on the incoming stream itself. DLD children may therefore have underspecified representations of real word stress patterns, leaving them at a disadvantage in analysing and learning new ones.

Evidence of a difficulty in discriminating acoustic cues to stress such as rise time, frequency and duration may result in a difficulty in processing the stress patterns of new words, leading to difficulties in developing robust templates of potential rhythmic templates in language, as well as difficulty in processing stress patterns as they occur 'online'. Accurate representations of phonological structure are thought to be necessary in order to achieve stable long-term representations (Gathercole & Baddeley, 1990). It seems likely that rhythmic structure could play a parallel role. If children with DLD have less robust long-term representations of rhythmic structure (Richards & Goswami, 2015), this suggests a difficulty in extracting rhythm from the signal, or underspecified rhythms, forming part of the 'noise' which could interfere with successful retrieval and consolidation into learning.

We are not aware of any previous studies which have specifically investigated rise time processing and NWR, however Goswami et al. (2002) included a one-syllable NWR task in their study of children with dyslexia. They found that rise time sensitivity accounted for 14% of unique variance in NWR once age, nonverbal IQ and vocabulary had been accounted for. In a later study (Corriveau, Pasquini, & Goswami, 2007), NWR was measured and combined with a word recall test to produce a combined working memory measure. Rise Time and Duration discrimination both contributed significant variance to working memory. Although it is not possible to state to what extent rise time contributed to NWR specifically in this study, it seems likely to have been a factor.

NWR tasks are therefore not merely proxy measures of PSTM, but are also influenced by levels of language knowledge. We consider that rhythm could be an under-explored linguistic factor influencing performance. There is some evidence for different performance in repeating short nonwords with different stress patterns in 2-year-old children. We wanted to explore this systematically across longer wordlengths in older children to see whether the intrinsic stress pattern of a nonword had a predictable effect on repetition accuracy. Children with DLD typically struggle with NWR tasks, and we consider that some of this difficulty may relate to the capacity to integrate stored and online representations of rhythm. We were therefore interested to discover whether the response pattern of children with DLD differed from that of TD children when stress patterns were manipulated. Finally, because of the potential role of rhythmic patterning, we thought that sensitivity to the acoustic cues of stress and rhythm, particularly rise time, frequency and duration, might be related to NWR task performance.

6.2 Devising the Task

Nonword repetition is a task in which children with DLD frequently encounter difficulty. Extrapolating from the results of Archibald & Gathercole (2007) we hypothesised that a difficulty in representing the underlying prosodic structure of the word could be contributing to the typically poor scores of children with DLD in this type of task.

Accordingly, we decided to examine the role of prosodic structure in detail in this task. Firstly, we wanted to know whether the specific stress pattern of the target stimuli would impact on the accuracy of children's responses; we therefore chose to systematically vary the location of the stressed syllable across wordlengths.

Furthermore, we wanted to know whether the stress status of an individual syllable (Stressed or unstressed) contributed to accuracy of repetition, and also, whether the position of a syllable relative to stress (i.e. preceding or following the stress in a word by a certain number of syllables) contributed to accuracy.

Finally, we hypothesised that listening to an entraining rhythm encapsulating the prosodic structure of the upcoming stimulus would support the children's representation of prosodic structure and thus assist them in performing the task. We hypothesised that by providing an external scaffold (rather than relying on an internally generated one) we would be able to boost the accuracy of their repetition.

We therefore set out to construct a stimulus set which would vary stress location systematically across wordlengths to create a range of target prosodic structures, for which we also created matching entraining rhythms.

6.3 Designing the Stimulus Materials

In constructing the nonword stimuli, we considered several structural features of potential words, such as the phonological content, syllable structure, wordlength and prosodic structure.

6.3.1 Phonological Content

Several aspects of phonology were considered in constructing the stimuli. These included the choice of consonants and vowels themselves and the combinations in which these occurred. These choices were motivated by consideration of phonological and articulatory development, frequency of occurrence and possible confounding factors such as existence as monosyllabic words.

6.3.1.1 Consonants

The Consonantal repertoire of children develops gradually over childhood with certain consonants (e.g. 'ch' /ʧ/) developing later than others (e.g. 'p'). We wanted to be sure that any repetition errors the children made were due to task factors rather than their stage of phonological development. We therefore took a conservative approach to consonant inclusion, considering only those which have typically been acquired by the age of 3 years 5 months (Dodd, Holm, Hua, & Crosbie, 2003). Since our participants were considerably older than this, it could be reasonably expected that these consonants would therefore form an established part of their phonological repertoire.

This first stage gave us a consonant pool of 17 consonants: /p, b, t, d, k, g, m, n, ŋ, f, v, s, z, h, w, l and j/.

The nasal plosive /ŋ/ was then discarded since it cannot occur in a syllable-initial position in English. /v/ and /z/ can occur word-initially in English, but are relatively infrequent in this position (Kessler & Treiman, 1997; Mines, Hanson, & Shoup, 1978) and so these were also discarded. My own clinical experience as a Speech and Language Therapist suggested that approximants such as /w, l, j/ and the glottal fricative /h/ can cause perceptual and production confusion at a later age than 3:5 and so it was decided to also discard these from the pool.

We were therefore left with a pool of ten consonants: /p, b, t, k, g, m, n, f/ and /s/

6.3.1.2 Vowels

One feature of the stress system in English is that vowels in unstressed syllables frequently occur in a reduced form as /ə/ (schwa). This therefore produces a phonological as well as an acoustic cue to the location of stress. Because we were interested in acoustics specifically, we therefore chose only vowels which can occur in their full form in unstressed syllables. From these parameters we looked at neighbourhood density and frequency data for CV rimes (De Cara & Goswami, 2003) which also tallied well with guides to overall phoneme frequency (Kessler & Treiman, 1997) and so chose three monophthong rimes and three diphthong rimes with the highest neighbourhood density and frequency. These were: monophthongs: ee /i/ as in ‘me’; or /ɔ/ as in ‘door’; oo /u/ as in ‘you’; diphthongs: ie /aɪ/ as in ‘pie’, ay /əɪ/ as in ‘pay’, ow /əʊ/ as in ‘low’.

6.3.2 Creating the CV Syllables

Combining the ten consonants with the six vocalic rimes created a pool of sixty potential CV syllables. Of these, forty-five were judged to be monosyllabic English words (e.g. poor, day, toe) and so were discarded. This also resulted in the loss of rime /ɔ/ and the consonants /b, s and m/ since all their potential combinations were real words. This left a total of fifteen CV syllables available for constructing the nonwords, from which twelve were chosen to give an even mix of monophthong and diphthong rimes.

We then constructed nonwords at lengths of 3, 4 and 5 syllables (Henceforth, Wordlength 3, Wordlength 4 and Wordlength 5). Each wordlength contained twelve words[16].

Throughout the stimuli, we observed several constraints on construction:

1) No phonemes were repeated within a word (consonant or rime)
2) No syllable occurred more than once in the same location for each Wordlength.
3) No CV syllable carried the stress more than once for each Wordlength.
4) Arrangement of monophthong and diphthong rimes were evenly distributed within each stress pattern.

6.3.3 Manipulating Stress

Each wordlength had twelve words. Stress patterns were systematically assigned in blocks of four, (e.g. Primary stress on the first syllable - 4 examples; Primary stress on second syllable – 4 examples; primary stress on third syllable – 4 examples).

For Wordlength 3, this meant that stress was distributed evenly across all three potential syllables (first, second, third). This was not possible for Wordlengths 4 and 5 without increasing the number of target words, which we felt would have made the task too onerous for the children to complete. We therefore used greatest frequency of occurrence according to the MRC Psycholinguistic Database[17] to determine the location of stress for these wordlengths[18]. For Wordlength 4, this resulted in stress falling on the 1st Syllable (i.e. Swww), 2nd syllable (wSww) and 3rd Syllable (wwSw)

---

[16] When referring to this task, ‘word’ here is used to indicate a ‘word-level unit’, which may be a nonword rather than a real word

[17] The MRC Psycholinguistic database is freely available at www.psy.uwa.edu.au/mrcdatabase/uwa_mrc.htm

[18] It should be noted that some words in the MRC database have been allocated two primary stressed syllables, but when frequencies were re-calculated discarding those marked with double stresses, a similar relative distribution obtained.

with frequency values of 31.9% (of all four-syllable words in the database), 42.7% and 37.9% respectively. For Wordlength 5, stress fell on the 2nd Syllable (wSwww), 3rd Syllable (wwSww) and 4th Syllable (wwwSw) – frequency values – 22%, 44% and 28% respectively.

6.4 Recording of the stimuli

Initially, tokens were recorded with each syllable falling at strictly 250ms intervals. However, on playback, these recordings were judged to be excessively unnatural and so the stimuli were re-recorded. This time the speaker had an entraining metronome beat playing in one ear (inaudible on the recording) in order to induce a beat at 4 Hz (250ms interbeat intervals) and produced the word aiming for the stressed syllable to fall in time with the beat. This resulted in a stimulus in which the inter-syllable interval approximated 250ms, but which did not sound unduly artificial.

An entrained version of each token was created by inserting an entraining rhythm prior to the spoken stimulus. The entraining rhythm consisted of a series of weak and strong beats following the pattern of the spoken stimulus, which was played twice prior to the spoken stimulus. In order to separate the two entraining rhythms, a ‘blank’ beat was inserted between each occurrence. This was the temporal equivalent of inserting an extra ‘phantom’ inter-stress interval, meaning that the second entraining rhythm remained consistent with the 750ms pulse.

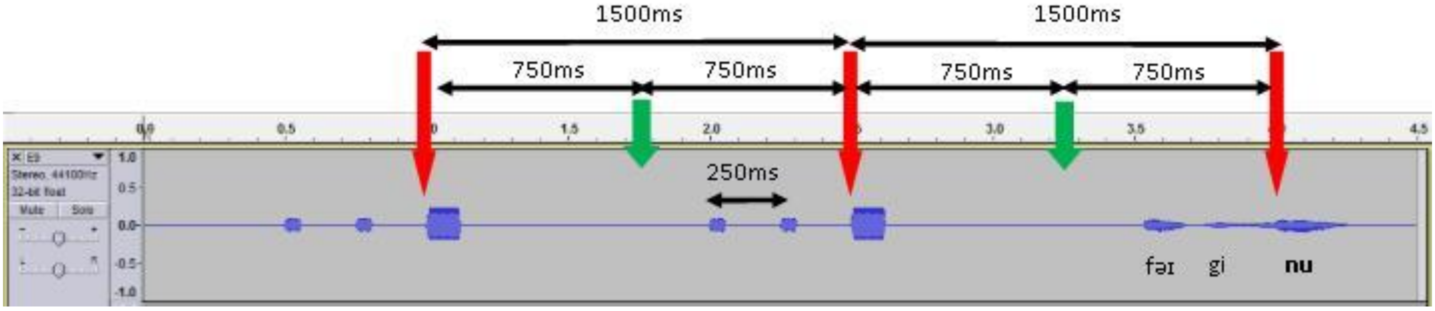


*Figure 6-1 Example soundwave of an entrained three-syllable stimulus, /fəɪgi**nu**/, with the stress on the third syllable. The red arrows indicate the location of strong beats. The green arrow indicates the location of the 'phantom beat'.*

The spoken token was aligned with the entraining rhythm so that the stressed syllable occurred 1500ms (i.e. 2 x 750ms) after the final strong beat of the entraining rhythm.

6.5 Results

The phonetic transcriptions of the children's responses were analysed at word-level (i.e. only a completely accurate repetition of a word scored as correct) and at syllable-level (i.e. a completely accurate repetition of that syllable scored as correct).

A sample (*n* = 21) of recordings was independently transcribed by a qualified Speech and Language Therapist who was blind to the overall purpose of the task and to participant group. Inter-rater agreement was calculated at word level (94.44%) and syllable level (94.38%).

Results were divided into words of three syllables (Wordlength 3), four syllables (Wordlength 4) and five syllables (Wordlength 5). Within each of these multiple levels, the effect of stress and the effect of the location of the stressed syllable was explored.

6.6 Overall Score

The maximum overall word score was 36, with a maximum word score for each Wordlength of 12. In order to compare across Wordlengths, syllable scores were calculated as a percentage of the target number of syllables correct.

Table 6-1 Results of one-way and repeated-measures ANOVAs for Word and Syllable scores for all words; and Wordlengths 3, 4 and 5.

Score	**AMC Mean (SD)**	**DLD Mean (SD)**	**YLC Mean (SD)**	***df***	***F***	***p***
Word Level – Number of Words Correct						
Overall Word Score[a,b,c] (max 36)	22.29 (5.56)	8.43 (4.65)	13.83 (7.1)	2	25.702	.000
Word Score Wordlength 3[a,c] (max 12)	10.67 (1.56)	5.5 (2.68)	7.48 (2.84)	2	23.061	.000
Word Score Wordlength 4[a,b,c] (max 12)	8.21 (2.47)	2.5 (1.5)	4.7 (2.95)	2	25.480	.000
Word Score Wordlength 5[a,c] (max 12)	3.42 (2.72)	0.43 (0.76)	1.65 (2.48)	2	7.913	.001
Syllable Level – Percentage of Syllables Correct						
Syllable Score Wordlength 3[a,c]	95.486 (5.604)	72.222 (14.122)	80.073 (16.383)	2, 58	17.162	.000
Syllable Score Wordlength 4[a,b,c]	86.198 (11.490)	54.762 (12.983)	67.935 (19.652)	2, 58	19.816	.000
Syllable Score Wordlength 5[a,c]	68.472 (17.067)	31.071 (12.483)	44.492 (24.957)	2, 58	17.921	.000

Note: a) AMC > DLD b) YLC > DLD c) AMC > YLC

In order to investigate differences between participant groups, a one-way ANOVA was conducted for each Wordlength at word and syllable level respectively. The AMC group scored significantly more highly than both DLD and YLC groups in each ANOVA (see Table 6-1). Post-hoc tests showed that the YLC group had a higher overall word score than the DLD group (*p* = .031), and their scores at word and syllable level for Wordlength 4 were significantly higher than the DLD group (see Table 6-1).

As expected, therefore, the DLD group scored less well than the AMC group throughout, replicating the results of many other studies of nonword repetition with this clinical group. The DLD children also tended to score less well than the YLC group, although this was only significant at Wordlength 4.

6.7 The effect of Wordlength on Accuracy of Words and Syllables

Previous studies have noted an effect of Wordlength on nonword repetition accuracy. We therefore examined if this was the case in our data using repeated-measures ANOVA (3 x 3 – Group [AMC, DLD, YLC] x Wordlength [Wordlength3, Wordlength4, Wordlength5]. This was conducted first with Word Score as the DV, then again with Syllable Score as the DV.

Table 6-2 Results of repeated-measures ANOVAs by Wordlength for Word and Syllable level accuracy

Score	**Wordlength 3 Mean (SD)**	**Wordlength 4 Mean (SD)**	**Wordlength 5 Mean (SD)**	***df***	***F***	***p***
Word Level – Number of Words Correct						
AMC[a, b, c]	10.67 (1.56)	8.21 (2.47)	3.42 (2.72)	2, 46	116.518	.000
DLD[a, b, c]	5.5 (2.68)	2.5 (1.5)	0.43 (.756)	1.314, 17.084	64.628	.000
YLC[a, b, c]	7.48 (2.84)	4.7 (2.95)	1.65 (2.479)	2, 44	64.102	.000
[19]All Groups[a, b, c]	8.28 (3.126)	5.57 (3.359)	2.07 (2.575)	2, 120	208.489	.000
Syllable Level – Percentage of Syllables Correct						
AMC [a, b, c]	95.486 (5.604)	86.198 (11.490)	68.472 (17.067)	1.478, 34.005	64.814	.000
DLD [a, b, c]	72.222 (14.122)	54.762 (12.983)	31.071 (12.483)	2, 26	133.266	.000
YLC [a, b, c]	80.073 (16.383)	67.935 (19.652)	44.492 (24.957)	2, 44	47.517	.000
All Groups,[a, b, c]	84.335 (15.64)	72.097 (19.622)	50.847 (24.592)	1.601, 96.052	170.260	.000

Note: a) Wordlength 3 > Wordlength; 4 b) Wordlength 3 > Wordlength 5; c) Wordlength 4 > Wordlength

There was a significant main effect of Wordlength for both Word Score, *F* (1.808, 104.879) = 204.249, *p* = .000, and Syllable score, *F* (1.645, 95.434) = 179.450, *p* = .000. In both cases, post-hoc

---

[19] The term 'All Groups' here and throughout the thesis refers to the results obtained when all participants are included in the analysis.

pairwise comparisons showed Wordlength 3 was significantly more accurate than Wordlengths 4 (*p* = .000) and 5 (*p* = .000). Wordlength 4 was also more accurate than Wordlength 5 (*p* = .000). The same pattern was found when the ANOVA was repeated for each group individually.

Accuracy of repetition therefore progressively decreased as Wordlength increased from Wordlength 3 down to Wordlength 5. This effect of Wordlength is clearly illustrated in Figure 6-2.

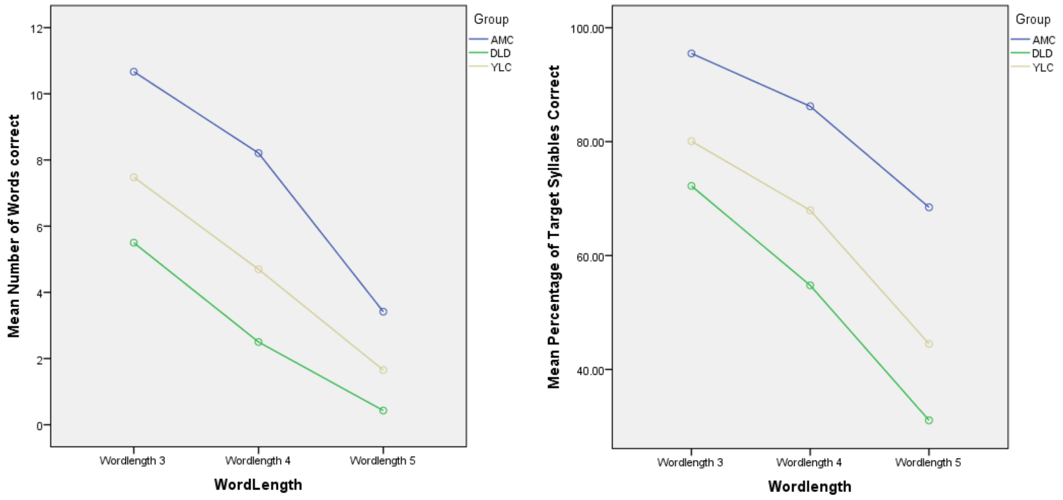


Figure 6-2 Graphs showing mean word and syllable score at each Wordlength by group

There were also significant Wordlength*Group interactions at both word (*F*(4) = 3.993, *p* = .005), and syllable scores (*F*(4) = 2.567, *p* = .042).

Wordlength therefore had a significant effect on the number of words and syllables repeated correctly for all groups, with the longer the word, the less accurate the repetition at both grain-sizes. Wordlength 3 was the most accurate, with significant decreases in accuracy for Wordlength 4 and then Wordlength 5. This result is consistent with previous work in nonword repetition.

6.8 Effect of the position of the Stressed Syllable

Each word at each Wordlength was then analysed according to the location of the stressed syllable. There were therefore three possible locations within each Wordlength. For three-syllable words, the stress could fall on the first syllable, (3:1 or Sww)[20] (e.g. **fu**təɪnaɪ, cf HOliday); second syllable (3:2 or wSw) (e.g. fəʊ**di**kaɪ, cf baNAna); or third syllable (3:3 or wwS) (e.g. naɪfu**təɪ**, cf magaZINE).

---

[20] In order to indicate prosodic structure of words, the format 'length:stress' will be used, so that, for example, a five-syllable word with the stress falling on syllable number three (wwSww) will be referred to as a word of type 5:3.

For four-syllable words, the stress could fall on the first (4:1 or Swww), second (4:2 or wSww) or third syllables (4:3 or wwSw). For five-syllable words, the stress could fall on the second (5:2 wSwww), third (5:3 or wwSww) or fourth syllables (5:4 or wwwSw). The maximum score for a single prosodic type (e.g. 3:2) was 4.

We predicted that the position of the stressed syllable within each Wordlength would affect repetition accuracy due to the different rhythmic patterns that this would create. As this was an exploratory study we had no firm predictions about the direction this would take.

In order to better present the pattern of the results, Wordlengths 3, 4 and 5 will be presented separately, with word- and syllable-level effects considered at each Wordlength.

6.8.1 Effect of Stressed Syllable Location - Wordlength 3

To investigate whether the location of the stressed syllable affected word-level accuracy, a repeated-measures ANOVA was conducted (3 x 3- Group[AMC, DLD, YLC] x Stress Location [3:1, 3:2, 3:3]. This was calculated firstly with Word Score as the DV, then Syllable Score as the DV.

At word level, there was a main effect of Stress Location F(1.770, 102.651) = 23.307, *p*= .000, with 3:3 structures more accurate than 3:2 and 3:1. In contrast, at syllable level there was no significant effect of Stress Location *F*(2,116) = 2.916, *p* = .058.

There was also a significant Stress Location * Group interaction *F*(4) = 4.329, *p* = .003, which Figure 6-3 suggests was due to the greater accuracy of 3:3 words for the DLD and the YLC groups.

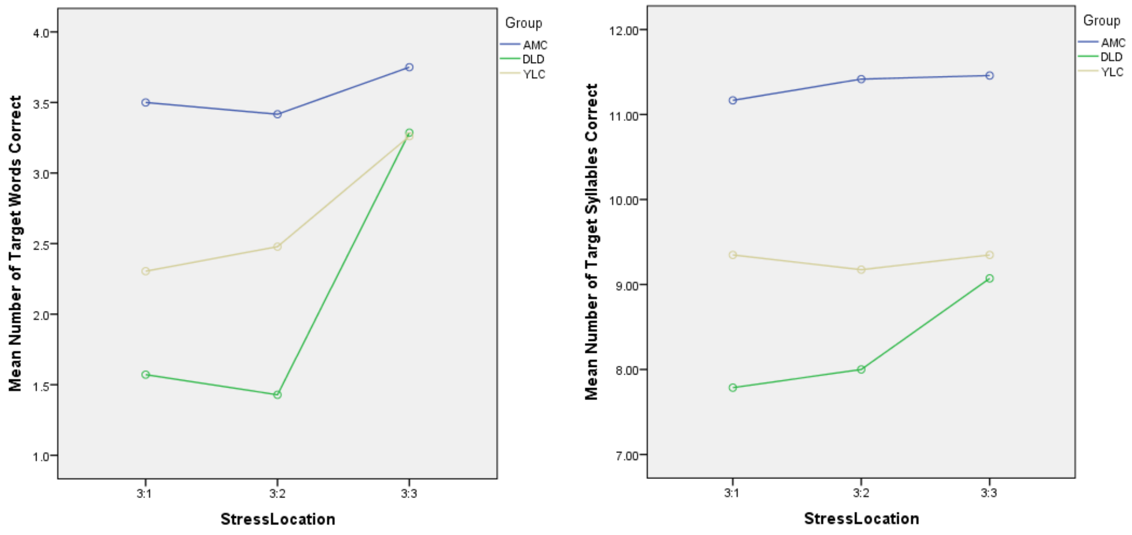


Figure 6-3 Graphs showing number of words and syllables correct by Stress Location at Wordlength 3

To explore the interaction, ANOVAs were therefore re-run for each group individually (1 x 3 – Group x Stress Location [3:1, 3:2, 3:3] – (Table 6-3).

Table 6-3 Number of words and syllables correct by group for 3:1, 3:2 and 3:3 words

Score	**Wordlength:Stress 3:1 (Sww)**	**Wordlength:Stress 3:2 (wSw)**	**Wordlength:Stress 3:3 (wwS)**	***df***	***F***	***p***
Word Score (max 4)						
AMC	3.5 (.933)	3.417 (.776)	3.75 (.532)	2, 46	1.565	.220
DLD[a, b]	1.571 (1.1016)	1.429 (.852)	3.286 (1.204)	2, 26	21.484	.000
YLC[a]	2.304 (1.146)	2.478 (1.31)	3.261 (1.01)	1.502, 33.046	4.985	.020
Syllable Score (max 12)						
AMC	11.167 (1.736)	11.417 (.776)	11.458 (.977)	2, 46	.507	.605
DLD[a]	7.786 (2.359)	8.00 (1.84)	9.0714 (2.056)	2, 26	4.174	.027
YLC	9.348 (2.269)	9.174 (2,887)	9.348 (2.145)	2, 44	.122	.885

Note: a) 3:3 > 3:1; b) 3:3 > 3:2

The individual ANOVAs (Table 6-3) revealed there was no effect of Stress Location for the AMC group at either word or syllable level. There was a significant effect for the DLD group with word type 3:3 more accurate than both than 3:2 (*p* = .000) and 3:1 (*p*= .002) at word level, 3:3 more accurate than 3:1 at syllable level (*p* = .012). 3:3 was also more accurate than 3:1 at word level for the YLC group (*p*= .035).

The position of the stressed syllable within the word therefore affects the accuracy of word-level repetition of three-syllable words. This effect is primarily driven by the DLD and YLC groups since there was no significant effect of stressed syllable position for the AMC group. For the DLD and YLC groups, repetition was significantly more accurate when repeating words in which the stress fell on the third, final syllable than when the stress fell on the first (DLD, YLC) or second (DLD) syllables. This effect was most marked when examining whole-word accuracy.

6.8.2 Effect of Stressed Syllable Location - Wordlength 4

Similar to the analysis of Wordlength 3, a repeated-measures ANOVA (3 x3 – Group [AMC, DLD, YLC] x StressLocation [4:1, 4:2, 4:3]) was conducted at both word and syllable levels.

There was a significant main effect of Stress Location at both word and syllable levels: Word: *F*(2, 116) = 10.775, *p* = .000, Syllable: Location *F*(2, 116) = 7.476, *p* = .001. At both levels, wordtypes 4:1 and 4:3 were more accurate than 4:2. There were also significant StressLocation * Group interactions – Word level: *F*(4) = 3.148, *p* = .017; Syllable level: *F*(4) = 4.098, *p* = .004.

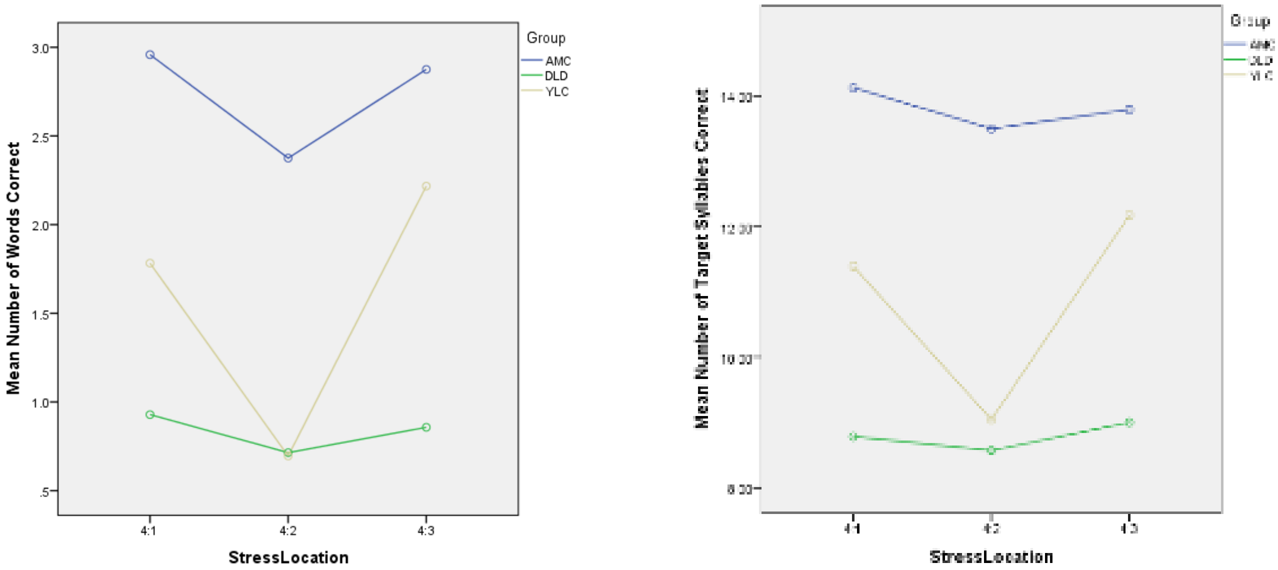


Figure 6-4 Graphs showing Number of Words and Syllables correct by Stress Location for Wordlength 4

Individual repeated-measures ANOVAs by Group (1 x 3 – Group x StressLocation [4:1, 4:2, 3:3] were conducted to explore these results further.

Table 6-4 Results of repeated-measures ANOVAs at Word and Syllable Level for Stress Locations at Wordlength 4

Score	**Wordlength:Stress 4:1 (Swww)**	**Wordlength:Stress 4:2 (wSww)**	**Wordlength:Stress 4:3 (wwSw)**	***df***	***F***	***p***
Word Score (max 4)						
AMC	2.96 (.999)	2.38 (1.279)	2.88 (1.116)	2, 46	2.582	.087
DLD	.93 (.616)	.71 (.914)	.86 (1.027)	2, 26	.220	.804
YLC^{a,b}	1.78 (1.313)	.70 (1.020)	2.22 (1.278)	2, 44	18.799	.000
Syllable Score (max 16)						
AMC	14.125 (1.963)	13.5 (2.28)	13.792 (2.587)	2, 46	.839	.439
DLD	8.786 (1.805)	8.571 (2.821)	9.0 (3.305)	2, 26	.139	.871
YLC^{a,b}	11.391 (3.258)	9.044 (3.937)	12.174 (3.366)	2, 44	15.684	.000

Note: a) 4:1 > 4:2; b) 4:3 > 4:2

We can see that the overall effect appears primarily driven by the YLC group who had significantly poorer scores for the 4:2 condition. The decrease in score for 4:2 approached significance for the AMC group at word level, but was not significant at syllable level. There was no effect of stress location for the DLD group.

Stress Location therefore had a significant effect on repetition accuracy at Wordlength 4. At this Wordlength, the effect was primarily driven by the YLC group, whose repetition of wordtype 4:2 (wSww) was significantly poorer than of 4:1 (Swww) and 4:3 (wwSw).

6.8.3 Effect of Stressed Syllable Location - Wordlength 5

The analyses were then repeated for Wordlength 5 with a repeated-measures ANOVA (3 x3 – Group [AMC, DLD, YLC] x StressLocation [5:2, 5:3, 5:4]) at both word and syllable levels.

There was no significant effect of Stress Location at word level, however the effect was significant at syllable level F(2, 116) = 5.819, p = .004. At syllable level, wordtypes 5:2 (wSwww) and 5:3 (wwSww) were more accurate than 5:4 (wwwSw) (*p* = .026, .016 respectively).

Table 6-5 Results of repeated-measures ANOVAs at Word and Syllable level for Stress Locations at Wordlength 5

Score	**Wordlength:Stress 5:2 (wSwww)**	**Wordlength:Stress 5:3 (wwSww)**	**Wordlength:Stress 5:4 (wwwSw)**	***df***	***F***	***p***
Word Score (max 4)						
AMC	.96 (.908)	1.5 (1.445)	.96 (.859)	2, 46	3.954	.026
DLD	.14 (.363)	.14 (.363)	.14 (.363)	2, 26	.000	1
YLC	.70 (1.020)	.57 (.896)	.39 (.941)	2, 44	1.581	.217
All Groups	.67 (.908)	.84 (1.20)	.56 (.866)	2, 166	2.128	.124
Syllable Score (max 20)						
AMC^{a,b}	13.667 (3.796)	14.833 (4.219)	12.583 (3.322)	2, 46	7.336	.002
DLD	6.643 (2.499)	6.286 (3.474)	5.714 (2.614)	2, 26	.952	.399
YLC	9.609 (4.727)	8.783 (5.559)	8.304 (5.489)	2, 44	2.319	.110
All Groups^{b,c}	10.525 (4.78)	10.590 (5.792)	9.393 (4.944)	2, 116	5.819	.004

Note: a) 5:3 > 5:2 b) 5:3 > 5:4; c) 5:2 > 5:4

Groupwise, there was a significant effect of position for the AMC group at both word and syllable levels. Pairwise comparisons were not significant at word level, but at syllable level, wordtype 5:3 was more accurate than 5:2 (*p* = .046) and 5:4 (*p* = .006) (AMC group).

Despite the AMC group being the only group showing a significant effect of Stress Location, there were no significant interactions between group and Stress Location at either word or syllable level. It should be borne in mind that scores at word level were low across all groups at this Wordlength.

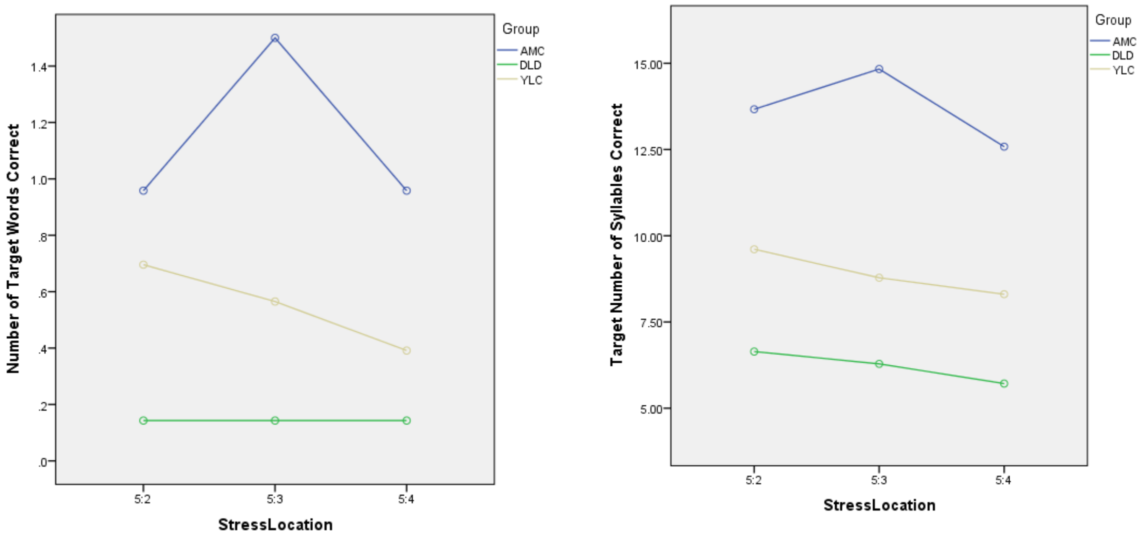


Figure 6-5 Graphs showing Number of Words and Syllables correct by Stress Location at Wordlength 5

Location of the stressed syllable therefore exerted a significant influence on repetition accuracy at Wordlength 5, however at this Wordlength this was significant only for the AMC group. For the AMC group, words with the structure 5:3 were more accurate than either 5:2 or 5:4.

6.8.4 Summary of Effect of Stress Location on Accuracy (Word and Syllable Level)

The location of a stressed syllable within a word has been shown to impact on how accurately that word is repeated. The effect of stressed syllable location, however, varied according to both Wordlength and participant group. For Wordlength 3, it was the DLD group which showed a significant effect of Stress Location with structures of type 3:3 (wwS) more accurate than 3:2 (wSw) and 3:1 (Sww). The YLC group also showed an effect at word level with 3:3 (wwS) again more accurate than 3:1 (Sww). For Wordlength 4, only the YLC group were affected by Stress Location, with 4:1 (Swww) and 4:3 (wwSw) structures more accurate than 4:2 (wSww). At Wordlength 5, the AMC group were the only group influenced by Stress Location, with 5:3 (wwSww) structures more accurate than both 5:2 (wSwww) and 5:4 (wwwSw).

6.9 Effect of Stress on Repetition Accuracy

Having established that the location of the stressed syllable within a word affects repetition accuracy, we investigated whether stress influenced repetition of individual syllables. We therefore calculated the percentage of target syllables repeated correctly, dividing targets into either stressed or unstressed syllables. We analysed this first across the entire dataset, and then examined effects for each Wordlength separately.

6.9.1 Effect of Stress on Repetition Accuracy – Entire dataset

The total percentage of target stressed and unstressed syllables repeated correctly was calculated and then paired sample t-tests were used to compare accuracy rates both for each group and across all groups.

Table 6-6 – Results of t-tests comparing accuracy of stressed and unstressed syllables

Score	**Stressed Mean (SD)**	**Unstressed Mean (SD)**	***df***	***t***	***p***
Percentage of Target Syllables Correct					
AMC[a]	87.612 (11.438)	78.974 (11.515)	23	9.255	.000
DLD[a]	58.929 (11.436)	46.032 (12.205)	13	10.522	.000
YLC[a]	70.048 (18.858)	58.253 (18.982)	22	8.858	.000
All groups[a]	74.408 (18.489)	63.601 (19.818)	60	15.210	.000

Note: a) Stressed > Unstressed

The results of the t-tests demonstrate that stress has a significant effect on syllable-level accuracy, with stressed syllables significantly more accurate than unstressed syllables for all groups (*p* = .000 throughout).

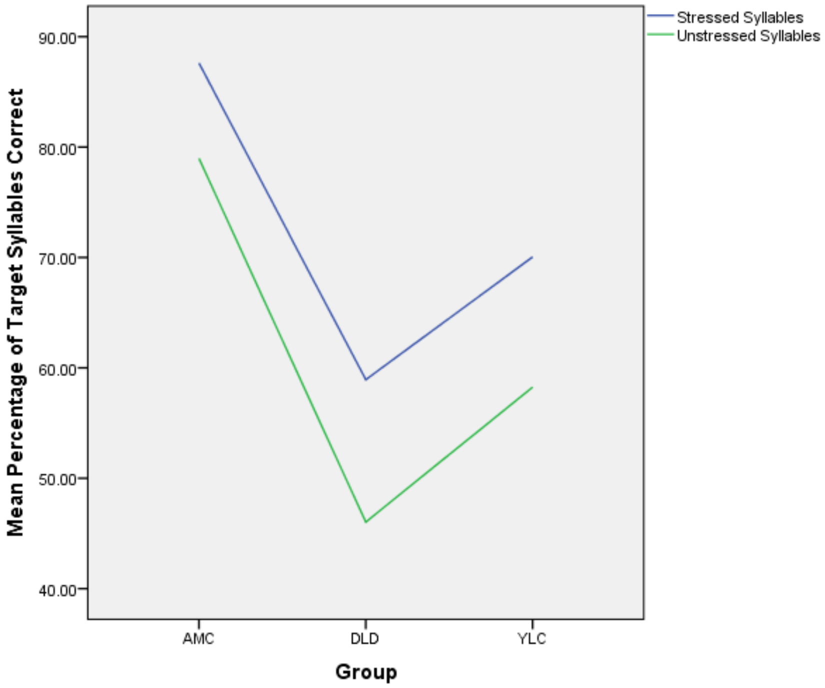


Figure 6-6 Graph showing Mean Percentage of Target Syllables Correct by Stress Status and Group

A repeated-measures ANOVA (3 x 2 – Group[AMC, DLD, YLC] x Stress [Stressed, Unstressed] revealed a significant main effect of stress *F*(1, 58) = 249.188, *p* = .000. There was also a significant interaction of group and stress *F*(2) = 3.443, *p* = .039. From Figure 6-6, we can see this arose because the effect of stress for the DLD and YLC groups was stronger than for the AMC group.

Stress therefore had a significant effect on accuracy across participant groups, with stressed syllables more accurate than unstressed syllables (*p* = .000 throughout). The effect of stress was particularly marked for the DLD and YLC groups, with greater differences between stressed and unstressed syllables for them than for the AMC group.

We then examined the differences between Stressed and Unstressed syllable accuracy at each Wordlength, again using paired sample t-tests.

Table 6-7 – Results of t-tests comparing stressed and unstressed syllables for each group at each Wordlength

Score	**Stress Mean (SD)**	**Unstressed Mean (SD)**	***df***	***t***	***p***
Wordlength 3					
AMC	96.181 (8.142)	93.880 (7.513)	23	2.020	.055
DLD[a]	76.786 (15.736)	66.518 (15.622)	13	5.641	.000
YLC	78.261 (18.247)	77.853 (17.166)	22	.258	.799
All groups[a]	84.973 (16.861)	81.557 (17.359)	60	3.502	.001
Wordlength 4					
AMC[a]	91.667 (11.261)	85.104 (12.253)	23	4.607	.000
DLD	57.738 (17.742)	55.714 (16.873)	13	.579	.573
YLC[a]	77.899 (19.399)	67.50 (19.803)	22	5.977	.000
All groups[a]	78.689 (20.611)	71.721 (20.032)	60	5.686	.000
Wordlength 5					
AMC[a]	91.435 (7.548)	74.708 (12.071)	23	10.850	.000
DLD[a]	78.175 (4.339)	46.152 (11.997)	13	12.763	.000
YLC[a]	83.213 (9.997)	56.383 (19.824)	22	10.971	.000
All groups[a]	85.291 (9.547)	61.245 (19.127)	60	16.432	.000

Note: a) Stress > Unstressed

As we saw with the effect of stress location, the effect of stress on repetition accuracy varied according to Wordlength and participant group. At Wordlength 3, the effect of stress was significant for the DLD group only. The inverse effect was found at Wordlength 4, with stress a significant influence for both AMC and YLC groups, whilst there was no effect for the DLD group. At Wordlength 5, stress was significant for all three groups.

6.10 Interaction of Stress with Wordlength

As the effect of stress varied according to Wordlength, we wondered whether the accuracy of stressed and unstressed syllables themselves varied significantly between Wordlengths. We therefore conducted a series of repeated-measures ANOVAS (3 x 3 - Group [AMC, DLD, YLC] x Wordlength [3, 4, 5] with first stressed, then unstressed syllables as the dependent variables (see Tables 6-8 and 6-9). These were also repeated for each group individually (1 x 3 - Group x Wordlength [3, 4, 5])

Table 6-8 – Results of repeated-measures ANOVAs for Stressed syllables by Wordlength

Score	**Wordlength 3 Mean (SD)**	**Wordlength 4 Mean (SD)**	**Wordlength 5**	***df***	***F***	***p***
Percentage of Stressed Syllables Correct						
AMC[b]	96.181 (8.142)	91.667 (11.261)	91.435 (7.548)	1.404, 32.289	5.424	.017
DLD[a, c]	76.786 (15.736)	57.738 (17.742)	78.175 (4.339)	2, 26	13.084	.000
YLC	78.261 (18.247)	77.899 (19.399)	83.213 (9.997)	2, 44	1.825	.173
All groups[a, c]	84.973 (16.861)	78.689 (20.611)	85.291 (9.547)	2, 120	8.004	.001

Note: a) Wordlength 3 > Wordlength 4; b) Wordlength 3 > Wordlength 5; c) Wordlength 5 > Wordlength 4

For stressed syllables, there was a significant effect of Wordlength for AMC and DLD groups, although the direction this took differed by group. For the AMC group, stressed syllables at Wordlength 3 were more accurate than Wordlength 5. In contrast, for the DLD group, stressed syllables at Wordlengths 3 and 5 were more accurate than stressed syllables at Wordlength 4. There was no effect of Wordlength for the YLC group.

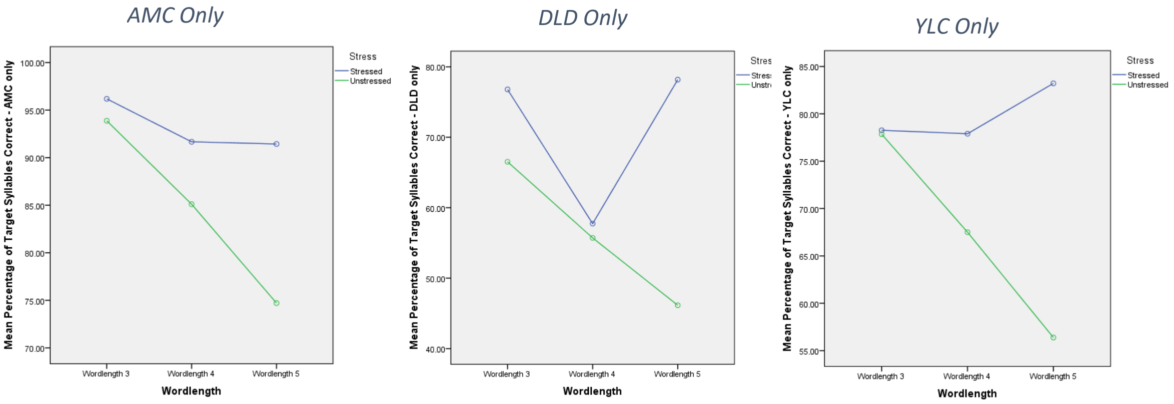


Figure 6-7 Graphs showing Mean Percentage of Target Syllables Correct for Stressed and Unstressed Syllables at Wordlengths 3, 4 and 5 for each group

From inspecting the graphs (Figure 6-7) we can clearly see the different response patterns affecting stressed syllables for the three participant groups.

Table 6-9 – Results of repeated-measures ANOVAS for Unstressed syllables by Wordlength

Unstressed Syllables	**Wordlength 3 Mean (SD)**	**Wordlength 4 Mean (SD)**	**Wordlength 5**	***df***	***F***	***p***
Percentage of Unstressed Syllables correct						
AMC[a,b, c]	93.880 (7.513)	85.104 (12.253)	74.708 (12.071)	2, 46	47.834	.000
DLD[a, b, c]	66.518 (15.622)	55.714 (16.873)	46.152 (11.997)	2, 36	25.742	.000
YLC[a, b, c]	77.853 (17.166)	67.50 (19.803)	56.383 (19.824)	2, 44	19.667	.000
All groups[a, b, c]	81.557 (17.359)	71.721 (20.032)	61.245 (19.127)	2, 120	74.551	.000

Note: a) Wordlength 3 > Wordlength 4; b) Wordlength 3 > Wordlength 5 c) Wordlength 4 > Wordlength 5

The pattern of response was different for unstressed syllables than for stressed syllables. For unstressed syllables, accuracy decreased steadily by Wordlength for all three participant groups, so that Wordlength 3 was the most accurate and Wordlength 5 the least.

This different effect of Wordlength for stressed and unstressed syllables was also found by a repeated-measures ANOVA (3 x 3 x 2 - Group [AMC, DLD, YLC] x Wordlength [3, 4, 5] x Stress [Stressed, Unstressed] – DV – Percentage of Syllables Correct). This showed a significant interaction of Wordlength*Stress $F(2, 116) = 111.094$, $p = .000$, as well as Wordlength*Stress*Group $F(4) = 9.867$, $p = .000$, supporting the previous figures that the effect of stress varied according to both Wordlength and group.

6.10.1 Summary of interaction of Stress with Wordlength

The effect of Stress status at different Wordlengths therefore took a different path for unstressed and stressed syllables. Unstressed syllables showed a steady decrease in accuracy as the Wordlength increased, and this was evident for all three participant groups. The situation for stressed syllables was more complex. The YLC group showed no Wordlength effect for stressed syllables, whilst the AMC group had a slightly higher rate of stressed syllable accuracy for Wordlength 3 within the context of high accuracy rates overall. The DLD group showed a different pattern again, with a significant decrease in stressed syllable accuracy at Wordlength 4, only for the stressed syllable accuracy rate to increase again at Wordlength 5.

6.11 Effect of Syllable Location on Accuracy

We explored the effect of syllable location on accuracy. By Syllable Location, we mean the location of occurrence of each individual syllable within a word, so that for the word ‘fəɪginu’; ‘fəɪ’ would be syllable 1, ‘gi’ syllable 2 and ‘nu’ syllable 3. We conducted a series of repeated-measures ANOVAs for each group at each Wordlength, with Syllable score as the dependent variable. (i.e. Wordlength 3: 1 x 3 – Group x Syllable [1, 2, 3] ; Wordlength 4: 1 x 4 – Group x Syllable [1, 2, 3, 4]; Wordlength 5: 1 x 3 – Group x Syllable [1, 2, 3, 4, 5].

Table 6-10 – Results of repeated-measures ANOVAs for Syllable Location by Wordlength and group

Syllable Score	**Syllable 1**	**Syllable 2**	**Syllable 3**	**Syllable 4**	**Syllable 5**	***df***	***F***	***p***
Wordlength 3								
AMC [c]	11 (1.103)	11.542 (.932)	11.5 (1.022)			2, 46	4.783	.013
DLD	8.643 (1.906)	8.00 (2.32)	8.214 (2.154)			2, 26	.818	.452
YLC	9.348 (2.479)	9.044 (2.567)	9.478 (2.064)			2, 44	.844	.437
All Groups	9.836 (2.115)	9.787 (2.471)	9.984 (2.172)			2, 116	.552	.577
Wordlength 4								
AMC	10.375 (1.469)	10.625 (1.789)	10.083 (2.062)	10.222 (1.606)		3, 69	.776	.511
DLD[f]	6.857 (2.345)	6.00 (2.253)	5.929 (1.730)	7.571 (2.138)		3, 39	3.004	.042
YLC	7.826 (2.462)	8.348 (2.587)	7.522 (3.16)	8.913 (2.521)		3, 66	3.763	.015
All Groups[f]	8.607 (2.538)	8.705 (2.830)	8.164 (2.956)	9.164 (2.339)		2.685, 155.753	5.489	.002
Wordlength 5								
AMC[a, b, d, e, I, j]	9.083 (1.792)	8.375 (2.356)	7.083 (2.669)	7.333 (2.259)	9.208 (2.553)	4, 92	14.053	.000
DLD[h, I, j]	3.713 (2.463)	2.929 (1.685)	2.786 (1.847)	3.714 (1.939)	5.5 (1.871)	4, 52	7.818	.000
YLC [f, g, h,I, j]	4.957 (2.513)	5.391 (3.448)	4.087 (3.204)	5.391 (3.13)	6.87 (3.709)	4, 88	13.027	.000
All Groups[a, d, f, g, h, I, j]	6.295 (3.196)	6.00 (3.421)	4.967 (3.23)	5.771 (2.895)	7.475 (3.249)	4, 232	27.380	.000

Note: a) **Syll1** > Syll3; b) **Syll1** > Syll4 c) **Syll2** > Syll1; d) **Syll2** > Syll3; e) **Syll2** > Syll4; f) **Syll4** > Syll3; g) **Syll5** > Syll1; h) **Syll5** > Syll2; i) **Syll5** > Syll3; j) **Syll5** > Syll4

The pattern of results for the individual Wordlengths can be most easily appreciated by looking at Figure 6-8.

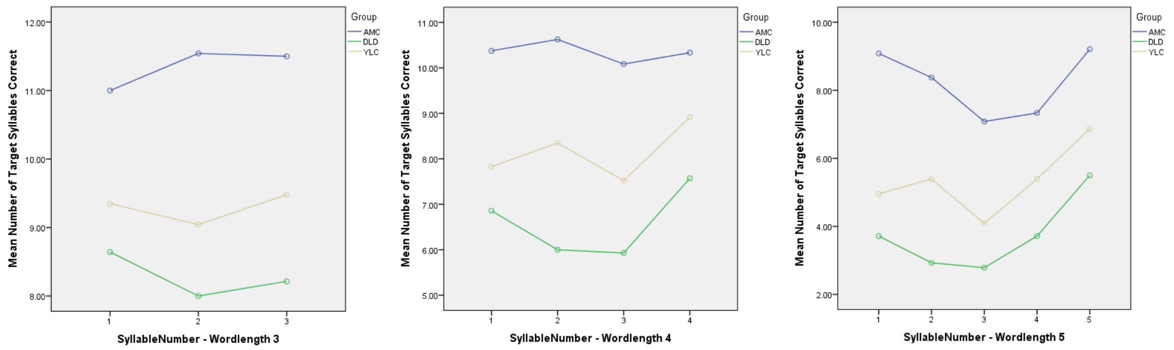


Figure 6-8 Graphs showing Mean Number of Target Syllables Correct by Syllable Location at Wordlengths 3, 4 and 5

There were limited effects of syllable location at Wordlengths 3 and 4. At Wordlength 3, only the AMC group showed a significant effect, with Syllable 2 more accurate than Syllable 1 (*p* = .019). At Wordlength 4, it was only the DLD group which showed an effect with Syllable 4 more accurate than Syllable 3 (*p* = .026). The YLC group showed a main overall effect of Syllable Location at Wordlength 4, but none of the pairwise comparisons were significant. In contrast, Syllable Location had a highly significant influence for all groups at Wordlength 5. The individual contrasts varied in detail (See Table 6-10 for details) however, Syllable 5 was the most accurate for all groups, with the AMC group also having Syllable 1 as relatively accurate.

Syllable Location therefore had a significant influence on repetition accuracy, with this effect marked for all groups at the longer Wordlength 5. From inspecting the graphs for the DLD group lines, we can see a similar trend to the DLD responses at Wordlength 4 that we see for all groups at Wordlength 5. This suggests that they are responding to similar influences, but at shorter Wordlengths, than the other two groups. The curve of responses at Wordlength 5 points towards greater accuracy for Syllable 1 and Syllable 5, i.e. for the first and final syllables of the word. This pattern would be consistent with primacy and recency effects which are found elsewhere in the short-term memory literature (e.g. Murdock, 1962), whereby the first and final items of lists tend to be better recalled. The results here indicate that primacy and recency can impact on accuracy levels within individual words as well as across word lists.

6.12 Interaction of Stress Syllable Position with Syllable Location

Within each Wordlength, we thought that syllable location may not tell the whole story, however, since we had also systematically varied the location of the stressed syllable. It could therefore be that the location of the stressed syllable, together with the syllable location, might provide a fuller account of accuracy distributions. We therefore used repeated-measures ANOVAS to study the interaction of syllable location and stressed syllable position at each Wordlength.

6.12.1 Stress Syllable Position and Syllable Location - Wordlength 3

We conducted a repeated-measures ANOVA (3 x 3 x 3 – Group [AMC, DLD, YLC] x Syllable Location [1, 2, 3] x Stress Location [3:1, 3:2, 3:3] with Syllable Score as the Dependent Variable, then ran the same analysis for each group individually. There was a significant interaction of Syllable Location and Stress Location, *F*(3.374, 195.673) = 6.597, *p* = .000. From the resulting graph, (Figure 6-9) we can see a pattern whereby the stressed syllable within each structure [3:1, 3:2, 3:3] is relatively accurate and this is true for all three possible syllable locations.

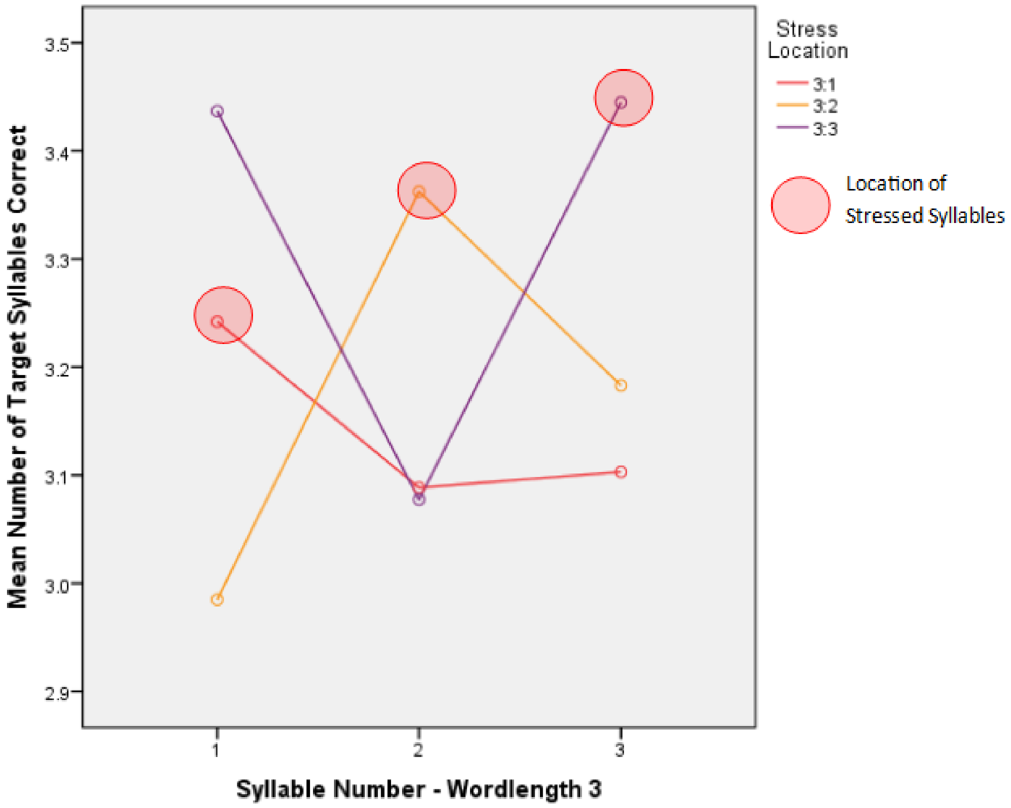


Figure 6-9 Graph showing the Mean Number of Target Syllables Correct by group together with stressed syllable locations for Wordlength 3

When repeated for each group individually, the interaction remained significant for the AMC group, *F*(2.414, 55.528) = 2.506, *p* = .047, and was particularly strong for the DLD group, *F*(4, 52) = 4.255, *p* = .005. There was no significant interaction for the YLC group, *F*(3.134, 68.953) = 1.125, *p* = .346.

Syllable Location therefore interacted with stress at Wordlength 3, with the stressed syllables showing a tendency towards greater accuracy. This overall effect was primarily driven by the DLD group.

6.12.2 Stress Syllable Position and Syllable Location - Wordlength 4

A similar repeated-measures ANOVA was conducted for Wordlength 4 (3 x 4 x 3 - Group [AMC, DLD, YLC] x Syllable Location [1, 2, 3, 4] x StressLocation [4:1, 4:2, 4:3] – DV Syllable Score). Again, there was a significant interaction of Syllable Location and Stress Location F(6, 348) = 18.142, *p* = .000, together with a significant interaction of Syllable Location, StressLocation and Group *F*(12) = 2.334 *p* = .007.

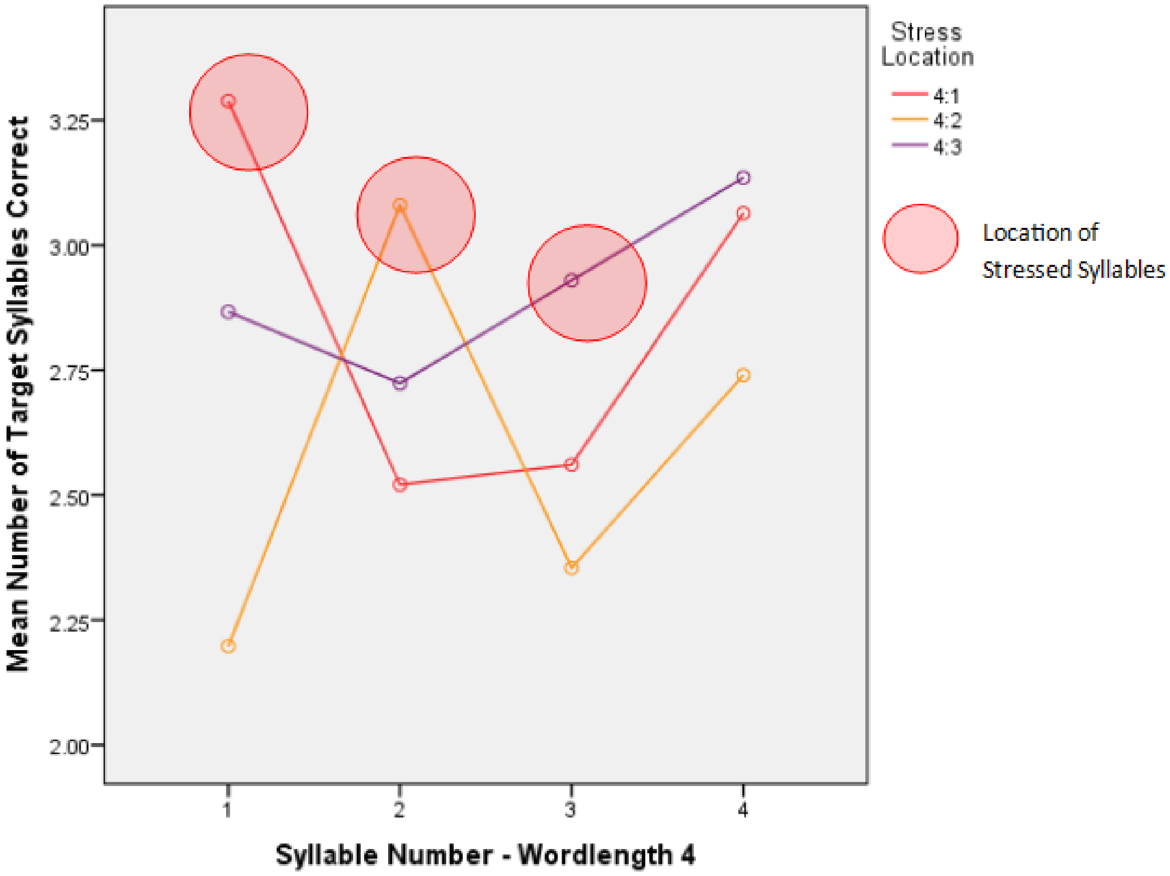
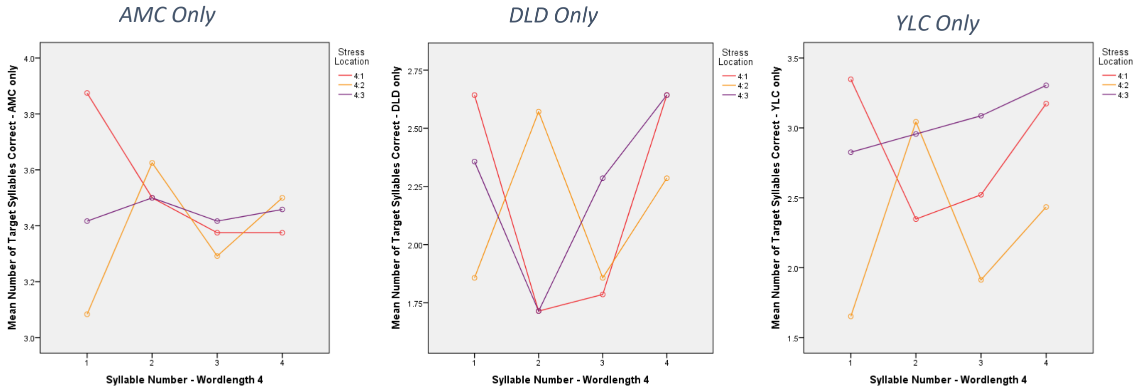


Figure 6-10 Graph showing the Mean Number of Target Syllables Correct for each syllable location together with the location of the stressed syllables for Wordlength 4

As at Wordlength 3, we can see a relatively privileged position for the stressed syllable across the syllable locations and stress locations. The interaction of stress with syllable location was also significant for each group individually: AMC: *F*(6, 138) = 4.729, *p* = .000; DLD *F* (6, 78) = 3.945, *p* = .002; YLC *F* (6, 132) = 12.651, *p* = .000.



AMC Only

DLD Only

YLC Only

Figure 6-11 Graphs for AMC, DLD and YLC groups showing Mean Number of Syllables Correct by Syllable Location for Wordlength 4

The Syllable Location * Stress Location * Group interaction is best interpreted by looking at the individual group profiles at this Wordlength (Figure 6-11). In doing so, we can see that the DLD group appears to be showing greater primacy and recency effects for structures 4:1 and 4:3 (red and purple lines) than the other two groups, whilst the pattern of effects for the 4:2 structure (yellow line) is similar across groups.

Stress location therefore also interacts with syllable location at Wordlength 4. At this Wordlength, the effect was highly significant for all three groups, with the stressed syllable tending towards greater accuracy across all prosodic structures.

6.12.3 Stress Syllable Position and Syllable Location - Wordlength 5

The repeated-measures ANOVA was conducted again at Wordlength 5: (3 x 5 x 3 - Group [AMC, DLD, YLC] x Syllable Location [1, 2, 3, 4, 5] x StressLocation [5:2. 5:3, 5:4] – DV Syllable Score). Again there was a significant interaction of Syllable Location and Stress Location *F*(6.333, 367.292) = 14.583, *p* = .000. The privilege for stressed syllables was less clear at this Wordlength, with the stressed third syllable of 5:3 structures less accurate than we might have expected compared with other syllable locations. As with Wordlength 4, the interaction of Syllable Location and Stress Location was significant for all three groups:AMC: *F*(5.073, 116.669) = 6.130, *p* = .000; DLD: *F*(8, 104) = 4.065, *p* = .000; YLC: *F*(8, 176) = 6.696, *p* = .000.

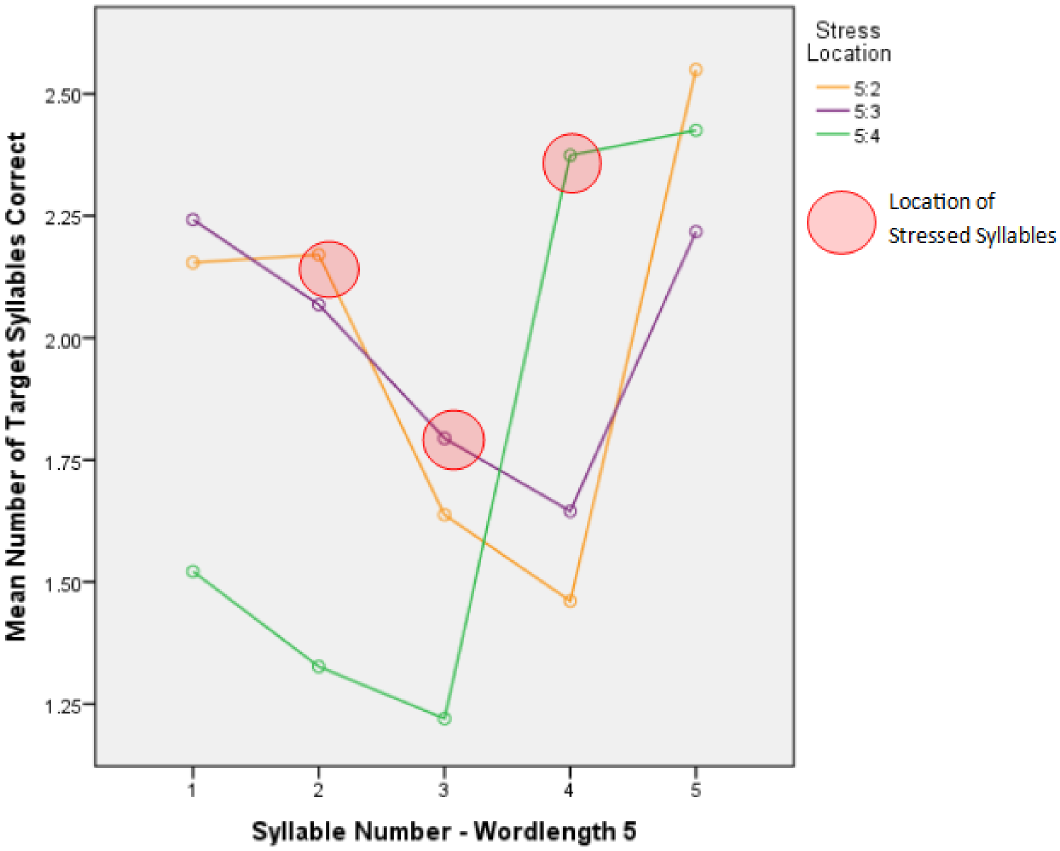


Figure 6-12 Graph showing the Mean Number of Target Syllables by Syllable Location together with the location of Stressed syllables for Wordlength 5

At this Wordlength, the apparent dip in accuracy for pre-stress syllables that we had observed in 3:2 and 4:2 structures is also not discernible. In contrast, the pattern appears to be one of general decline in accuracy towards the middle of the word (syllable 3) with better accuracy at the beginning and end of the words. This word-edge accuracy is then supported by the presence of stress (5:2 and 5:4), particularly noticeable in the 5:4 structure, which has much poorer accuracy at the beginning of the word, followed by a leap in accuracy at the stressed syllable. It seems, however that the central, stressed syllable of 5:3 structures is not sufficient to over-ride the negative effects of its word-central position.

6.12.4 Summary of interaction of stress and syllable location

The influence of stress therefore interacts with syllable location across Wordlengths. There appear to be two competing influences on syllable repetition accuracy – the stress level of syllables, with the stressed syllable tending towards greater accuracy, and the position of a syllable within a word – particularly whether it is word-initial, word-medial or word-final. At Wordlength 3, we saw that the stressed syllable was particularly influential for the DLD group, whilst the interaction between stress and syllable location was significant for all groups at Wordlengths 4 and 5. It seems that by Wordlength 5, it is syllable location which has the greater influence. Concurrently, a stressed syllable falling at word-edges provides a supportive effect for accurate repetition, however the privilege of stress is insufficient to offset the negative effects of a word-central syllable position. For the DLD group, this effect begins to be discernible at Wordlength 4, with their accuracy patterns for 4:1 and 4:3 structures mirroring the later pattern for 5:2 and 5:4 structures. The previously observed phenomenon of an accuracy dip for pre-stress syllables was seen in 3:2 , 3:3, and 4:2 structures, but not in 5:2, 5:3 or 5:4 structures. The influence of stress on repetition therefore alters depending on both the prosodic structure of the target word, and its Wordlength.

6.13 Position of Syllables relative to Stress

To explore the influence of a syllable’s position relative to stress further (e.g. pre-stress, post-stress etc), syllables were assigned a position according to whether they preceded (indicated by -), or followed (+) the stressed syllable, and by how many syllables. For example, for the 4:2 word ‘naɪ**təɪ**kudi’, ‘naɪ’ would be -1, ‘təɪ’ stress, ‘ku’ +1 and ‘di’ +2. The influence of a syllable’s position relative to stress was examined at each Wordlength.

6.13.1 Position of Syllables relative to Stress - Wordlength 3

A repeated-measures ANOVA (3 x 5 – Group [AMC, DLD, YLC] x Syllable position relative to Stress [-2,-1, STRESS, +1, +2) , DV – Percentage of syllables correct) was conducted across the dataset and for each group individually.

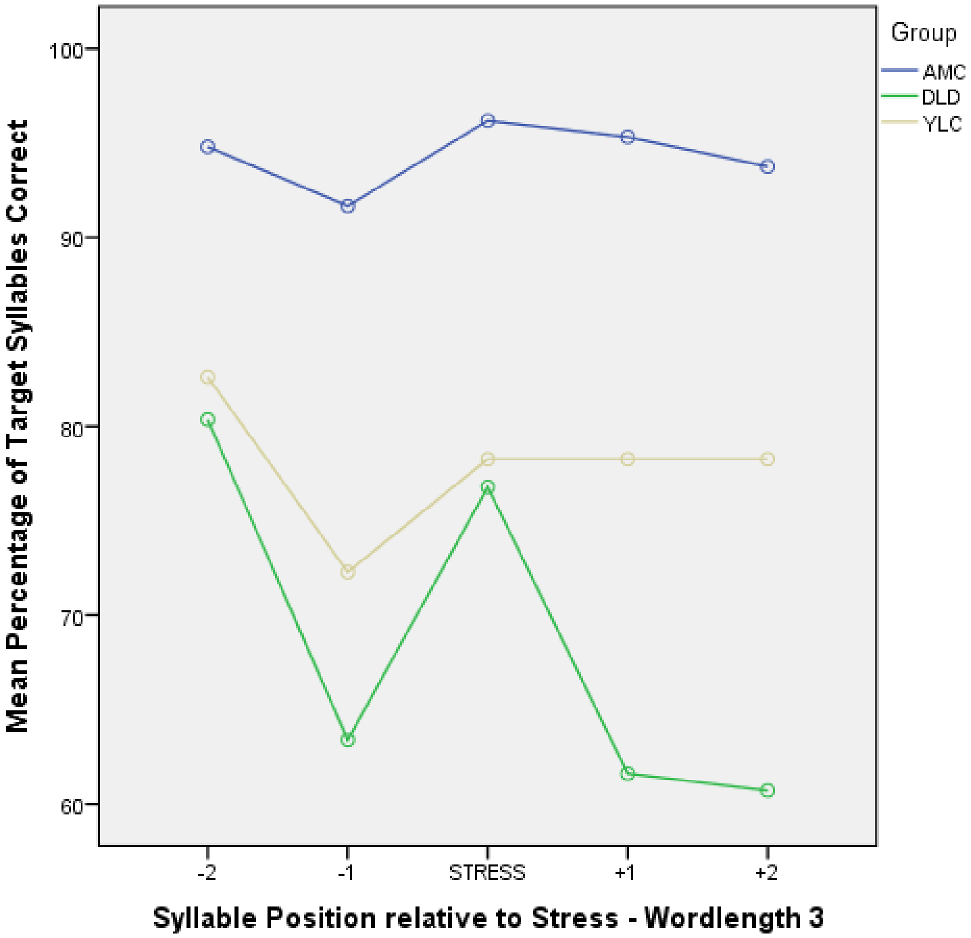


Figure 6-13 Graph showing Mean Percentage of Target Syllables Correct by Syllable Position relative to Stress for Wordlength 3

There was a significant main effect of syllable position *F*(3.107, 180.216) = 6.320, *p* = .000, with the stressed syllable more accurate than the syllables occurring either side of it: -1 (*p* = .000) and +1 (*p* = .008). The -2 syllable was also more accurate than -1. Due to the nature of the task construction, the -2 syllable would have corresponded to the first syllable of a 3:3 structure word (e.g. fəɪ of fəɪgi**nu**).

Table 6-11 Results of repeated-measures ANOVAS for each group for Syllable Position Relative to Stress

Score	**Relative Position -2**	**Relative Position -1**	**Relative Position STRESS**	**Relative Position +1**	**Relative Position +2**	***df***	***F***	***p***
Percent of Syllables Correct Wordlength 3								
AMC	94.79 (10.371)	91.667 (10.851)	96.181 (8.142)	95.313 (8.887)	93.75 (15.198)	2.107, 48.462	.864	.433
DLD[a, c]	80.36 (17.482)	63.393 (19.281)	76.786 (15.736)	61.607 (18.647)	60.71 (25.409)	2.677, 34.801	6.572	.000
YLC	82.61 (17.573)	72.283 (28.193)	78.261 (18.247)	78.261 (21.392)	78.26 (24.197)	2.875, 63.245	1.227	.307
All Groups[a, b, d]	86.89 (16.206)	77.869 (23.54)	84.973 (16.861)	81.148 (21.118)	80.33 (24.628)	3.107, 180.216	6.320	.000

Note: a) ***STRESS*** *> -1; b)* ***STRESS*** *> +1; c)* ***STRESS*** *> +2; d)* ***-2*** *> -1*

There was also a significant Syllable Position * Group interaction, F(8) = 2.281, p = .023. From Figure 6-13, this appears due to the effect of relative stress position being greater for the DLD group than for the other groups, coupled with a greater tail-off in accuracy for +1 and +2 syllables for this group. This was explored further using one-way repeated-measures ANOVAs for individual groups exploring Syllable Position [-2, -1, STRESS, +1, +2).

These revealed a significant effect of Relative Syllable Position for the DLD group *F*(2.677, 34.801) = 6.572, *p* = .000, but no significant effect for the AMC (*p* = .433) or YLC (*p* = .307) groups.

The position of a syllable relative to stress was therefore a significant factor in accuracy at Wordlength 3, with the stressed syllable more accurate than the syllables either side of it (-1 and +1). This effect was driven by the DLD group, who showed greater rises and falls in accuracy for the different syllable positions.

6.13.2 Position of Syllables relative to Stress - Wordlength 4

A repeated-measures ANOVA (3 x 5 – Group [AMC, DLD, YLC] x Syllable Position relative to Stress [-2, -1, STRESS, +1, +2, +3] - DV Percentage of Syllables Correct) was conducted across the dataset and then for each group individually.

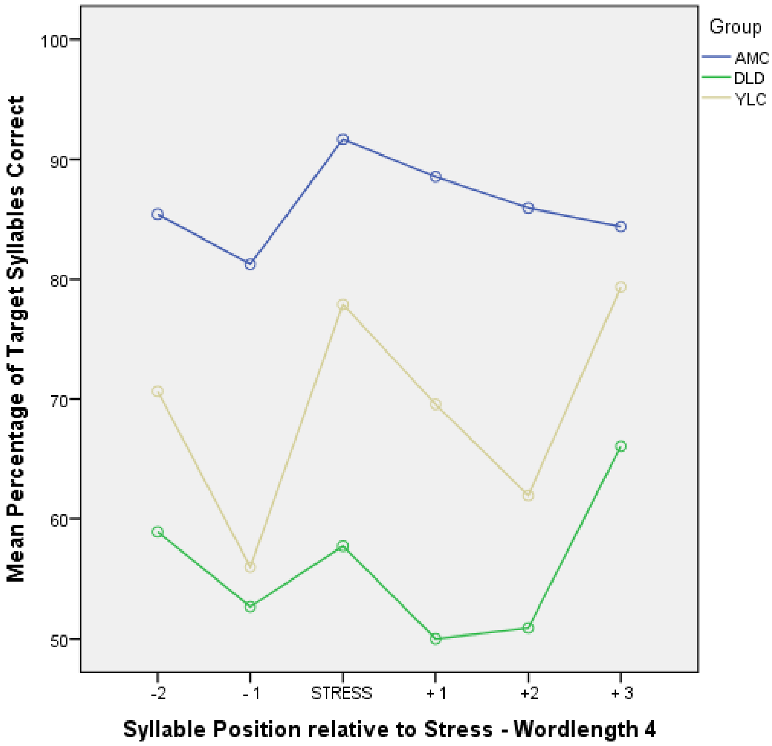


Figure 6-14 Graph showing Mean Percentage of Syllables Correct by Syllable Position relative to Stress at Wordlength 4

In the main ANOVA, there was a significant main effect of Position relative to Stress, *F*(2.627, 152.346) = 7.026, *p* = .000, with the Stressed syllable more accurate than -1, +1 and +2 and the syllable at +3 more accurate than -2 and +2. The syllable at -2 was also more accurate than -1.This represents an overall pattern to the results in which the stressed syllable and those at the beginning and end of the word (reflected in the -2 and +3 positions) are more accurate than others. This pattern is most easily discerned by looking at Figure 6-14.

There was a significant Position*Group interaction *F*(10) = 2.231, *p* = .016. From Figure 6-14, this appears to have arisen because the effect of stress, word-initial, word-final syllable positions appear to be greater for the YLC group. This interaction was further explored by running the repeated-measures ANOVA again each group individually (See Table 6-12).

Table 6-12 Results of repeated-measures ANOVAs for each group for Syllable Position Relative to Stress

Score	**Relative Position -2**	**Relative Position -1**	**Relative Position STRESS**	**Relative Position +1**	**Relative Position +2**	**Relative Position +3**	***df***	***F***	***p***
Percentage of Syllables Correct Wordlength 4									
AMC	85.42 (24.358)	81.25 (18.058)	91.667 (11.261)	88.542 (14.91)	85.938 (14.888)	84.38 (19.242)	1.913, 44.008	1.498	.235
DLD[b]	58.93 (28.768)	52.679 (23.09)	57.738 (17.742)	50 (18.777)	50.893 (17.991)	66.07 (25.205)	2.768, 35.979	1.873	.155
YLC[b, d, e, f, g, h]	70.65 (27.85)	55.978 (22.572)	77.899 (19.399)	69.565 (21.557)	61.957 (27.041)	79.35 (22.172)	2.778, 61.107	8.774	.000
All Groups[a, c, d, f, g, h]	73.77 (28.294)	65.164 (24.486)	78.689 (20.611)	72.541 (23.587)	68.852 (25.169)	78.28 (22.581)	2.627, 152.346	7.026	.000

Note: a) **STRESS** > -2; b) **STRESS** > -1; c) **STRESS** > +1; d) **STRESS** > +2; e) **+1** > -1; f) **+3** > -1; g) **+3** > +2; h) **-2** > -1

At this Wordlength, there was also therefore a significant effect of syllable position relative to stress. Whilst at Wordlength 3 the overall effect was primarily driven by the DLD group, at Wordlength 4 the effect was driven by the YLC group. Again, the pattern was of greater accuracy for the Stressed syllable than those surrounding it (-1 and +1) together with greater accuracy for syllables occurring word-initially (-2) and word-finally (+3).

6.13.3 Position of Syllables relative to Stress - Wordlength 5

A further repeated-measures ANOVA (3 x 7 – Group [AMC, DLD, YLC] x Syllable Position relative to Stress [-3, -2, -1, STRESS, +1, +2, +3] – DV Percentage of Syllables Correct) was conducted, together with analyses for each group individually.

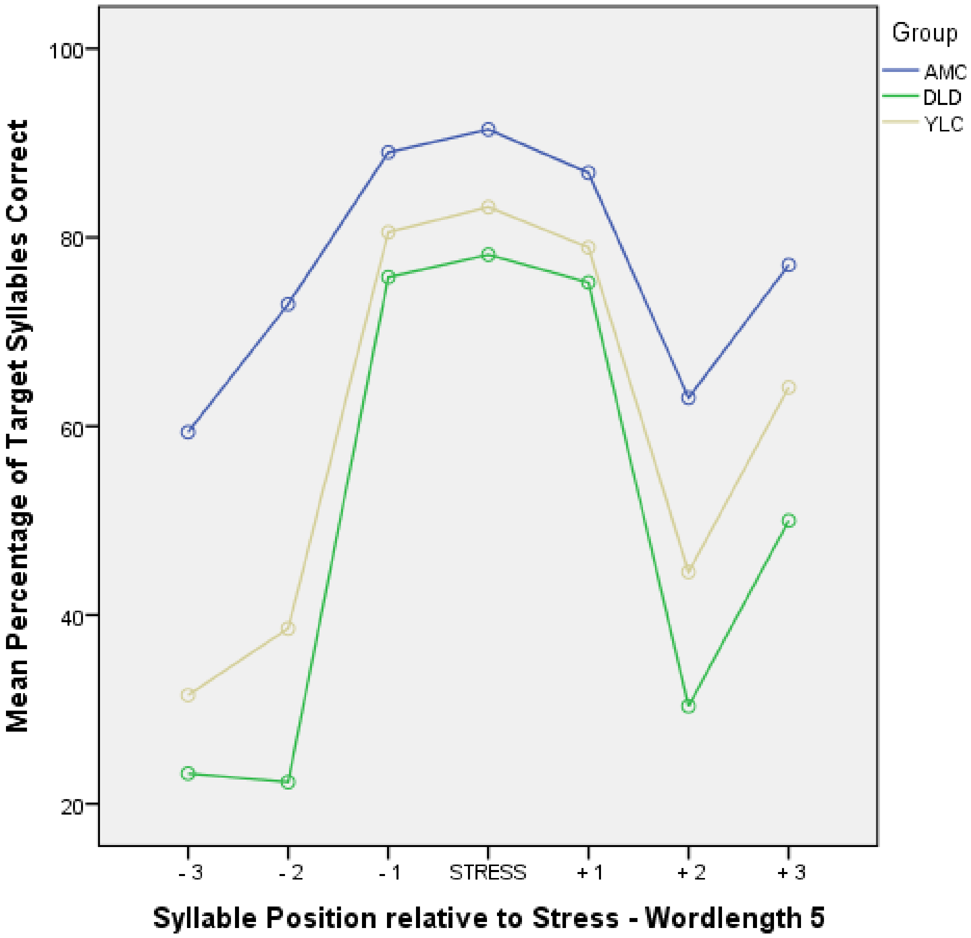


Figure 6-15 Graph showing Mean Percentage of Target Syllables Correct for each Syllable Position relative to stress by group for Wordlength 5

There was a significant main effect of Syllable Position relative to Stress, *F*(3.486, 202.179) = 100.888, *p* = .000. Post-hoc tests revealed many significant pairwise comparisons (see Table 6-13 for details) with the pattern being one of greater accuracy for the Stress syllable, those either side of it (-1 and +1) and for the +3 syllable, reflecting its position at the end of a word. This is most easily perceived in Figure 6-15.

There was also a significant interaction of Syllable Position and Group *F*(12) = 5,219, *p* = .000. From Figure 6-15, this appears to be due to the greater impact of Syllable Position relative to Stress for the DLD and YLC groups.

Table 6-13 Results of repeated-measures ANOVAs for each group for Relative Syllable Position to Stress at Wordlength 5

Score	**Relative Position -3**	**Relative Position -2**	**Relative Position -1**	**Relative Position STRESS**	**Relative Position +1**	**Relative Position +2**	**Relative Position +3**	***df***	***F***	***p***
Percent of Syllables Correct Wordlength 5										
AMC[a, b, c, d, e, f, g, h, j, k, l]	59.38 (21.885)	72.917 (19.736)	89.005 (5.82)	91.325 (7.548)	86.849 (8.79)	63.021 (20.35)	77.08 (24.358)	3.137, 72.143	20.085	.000
DLD[a, b, d, e, f, g, h, i, j, k, l, m, o]	23.21 (30.167)	22.321 (17.11)	75.794 (5.274)	78.175 (4.339)	75.223 (5.122)	30.357 (18.157)	50.00 (25.944)	2.869, 37.297	37.683	.000
YLC[a, b, c, d, f, g, h, j, k, l, n, o, p]	31.52 (26.347)	38.587 (29.415)	80.556 (7.397)	83.213 (9.997)	78.94 (9.573)	44.565 (29.394)	64.13 (33.563)	3.050, 67.092	44.927	.000
All Groups[a, b, d, e, f, g, h, i, j, k, l, m, n, o, p]	40.57 (29.638)	48.361 (31.08)	82.787 (8.237)	85.291 (9.547)	81.199 (9.589)	48.566 (26.722)	65.98 (29.942)	3.486, 202.179	100.888	.000

Note: a) **STRESS** > -3; b) **STRESS** > -2; c) **STRESS** > +1; d) **STRESS** > +2; e) **STRESS** > +3; f) **-1** > -3; g) **-1** > -2; h) **-1** > +2; i) **-1** > +3; j) **+1** > -3; k) **+1** > -2; l) **+1** > +2; m) **+1** > +3; n) **+3** > -3; o) **+3** > -2; p) **+3** > +2.

When conducted individually by group, the main effect of Syllable Position relative to Stress was significant for all groups: AMC: *F*(3.137, 72.143) = 20.085, *p*= .000; DLD: *F*(2.869, 37.297) = 37.683, *p* = .000; YLC: *F*(3.050, 67.092) = 44.927, *p* = .000. There were minor individual differences in the significance of pairwise comparisons (See Table 6-13 for details), however the overall pattern of significantly greater accuracy for syllables -1, STRESS, +1 and +3 above -3, -2 and +2 syllables was similar across groups.

Syllable position relative to Stress was therefore a significant factor in accuracy at Wordlength 5. This was the case for all groups, with the Stress syllable and its surrounding syllables (-1, +1) more accurate than other syllables. There was also increased accuracy for the +3 syllable, which always occurred in a word-final position. The pattern of results at this Wordlength therefore differed from Wordlenths 3 and 4. At shorter Wordlengths, the -1 and +1 syllables had undergone a dip in accuracy relative to the Stressed syllable, however at this Wordlength it was the Stress together with -1 and +1 which had the greater accuracy. This suggests that at longer Wordlengths, proximity to the stressed syllable has a supportive effect on accuracy.

6.13.4 Summary of effect of Syllable Position relative to Stress

The position of a syllable relative to the stressed syllable influences the accuracy of its repetition. The impact of this varies according to participant group and Wordlength. At Wordlength 3, the effect was greatest for the DLD group, with AMC and YLC groups showing no overall effect. Conversely, at Wordlength 4, it was the YLC group which showed a significant effect. At Wordlength 5, however, the effect of syllable position relative to stress was significant for all groups.

The overall pattern of results indicates a particular privilege of accuracy for the Stressed syllable, and for those syllables which occur towards the beginning and end of a word. This suggests that accuracy is affected by two complementary influences – one being the stress level, and the second being syllable location in terms of the primacy and recency effect documented in Section 6.11. The effects of stress at Wordlength 5 are particularly interesting, since they suggest that rather than a dip in accuracy for surrounding syllables as observed at Wordlengths 3 and 4, there is in fact a preservation effect surrounding the stressed syllable at this longer Wordlength.

6.14 Relationship of Nonword Repetition to Acoustic Thresholds

We hypothesised that scores on nonword repetition would relate to the children’s sensitivity to changes in the acoustic markers of stress, namely rise time, frequency, duration and intensity. We predicted that the greater the children’s sensitivity, the more accurate their repetitions would be. To discover if this was the case, we conducted correlations between the four acoustic measures (See Chapter 4) and nonword repetition accuracy at word and syllable level.

Table 6-14 Pearson correlations (r) one-tailed for AMC and DLD groups for Acoustic Processing Thresholds and Nonword Repetition Scores

Task	Duration	Frequency	Intensity	Word Score	Syllable Score
Rise Time	.340*	.795***	-.227	-.467**	-.514***
Duration	-	.406**	.016	-.468**	-.457**
Frequency	-	-	-.206	-.612***	-.673***
Intensity	-	-	-	.061	.032

*Note: * p < .05; ** p < .01; ***p < .001*

There was a significant correlation between Rise Time thresholds and scores on Nonword Repetition (Word: *p* = .002; Syllable: *p* = .000). Nonword Repetition accuracy also correlated with Duration (Word: *p* = .002; Syllable: *p* = .002) and Frequency thresholds (Word: *p* = .000; Syllable: *p* = .000).The inverse relationship indicates that the smaller the threshold value, the better the score on the task i.e. that children who were able to distinguish between smaller differences in Rise Time, Frequency and Duration tended to be more accurate in Nonword Repetition. There was no significant relationship with Intensity thresholds.

Better acoustic processing, as measured by smaller thresholds in the acoustic tasks of Rise Time, Duration and Frequency, was therefore related to more accurate Nonword Repetition.

In order to further explore the unique contribution of each variable to individual differences in performance, a series of regressions were conducted using the predictors of Age, NVIQ and each AT measure in turn. Step 1 was always entered as Age (months), Step 2 as NVIQ and Step 3 as the AT measure. Separate analyses were conducted with Word Score and then Syllable Score as the dependent variables.

Table 6-15 Results of Regressions exploring the unique variance in Word Score accounted for by Age (months), NVIQ and AT measures (AMC and DLD groups)

Word Score	**b**	**SEb**	**β**	***t***	***p***	**ΔR²**	***p***
Model 1							
Age	.074	.077	.146	.962	.343	.101	.055
NVIQ	1.248	.592	.320	2.109	.043	.154	.012
Rise Time	-.031	.016	-.307	-1.951	.060	.077	.060
Model 2							
Age	.080	.073	.158	1.100	.279	.101	.055
NVIQ	1.201	.567	.308	2.117	.042	.154	.012
Duration	-.062	.024	-.374	-2.546	.016	.122	.016
Model 3							
Age	.033	.081	.063	.410	.685	.102	.062
NVIQ	.999	.668	.244	1.496	.145	.167	.011
Frequency	-6.279	2.377	-.445	-2.641	.013	.134	.013
Model 4							
Age	.112	.079	.226	1.414	.167	.095	.068
NVIQ	1.530	.597	.402	2.561	.015	.155	.013
Intensity	-.082	.861	-.015	-.095	.925	.000	.925

Since Frequency, Duration and Rise Time all added unique variance (13.4% to 7.7%) when added at Step 3, a further model was created to further examine the influence of each AT variable. Frequency (with the greatest amount of unique variance) was added at Step 3, followed by Duration and then Rise Time.

Table 6-16 Results of Regressions exploring the unique variance in Word Score accounted for by Age (months), NVIQ and AT measures (AMC and DLD groups)

Word Score	**b**	**SEb**	**β**	***t***	***p***	**ΔR²**	***p***
Model 5							
Age	.029	.079	.055	.371	.714	.102	.062
NVIQ	.859	.654	.210	1.314	.199	.167	.011
Frequency	-4.160	3.630	-.295	-1.146	.261	.134	.013
Duration	-.051	.026	-.295	-1.934	.063	.067	.060
Rise Time	-.005	.024	-.049	-.216	.830	.001	.830

Table 6-17 Results of Regressions exploring the unique variance in Syllable Score accounted for by Age (months), NVIQ and AT measures (AMC and DLD groups)

SyllableScore	b	SEb	β	*t*	*p*	ΔR^2	*p*
Model 1							
Age	.126	.168	.112	.753	.457	.092	.068
NVIQ	2.704	1.291	.310	2.094	.044	.160	.011
Rise Time	-.083	.034	-.372	-2.425	.021	.113	.021
Model 2							
Age	.159	.163	.141	.980	.334	.092	.068
NVIQ	2.747	1.272	.315	2.159	.038	.160	.011
Duration	-.140	.055	-.376	-2.556	-.015	.124	.015
Model 3							
Age	.048	.174	.040	.274	.786	.101	.063
NVIQ	1.746	1.428	.190	1.223	.231	.157	.014
Frequency	-17.096	5.083	-.540	-3.364	.002	.198	.002
Model 4							
Age	.242	.178	.218	1.361	.183	.086	.083
NVIQ	3.475	1.346	.406	2.583	.015	.160	.012
Intensity	-.662	1.941	-.053	-.341	.735	.003	.735

Since Frequency, Duration and Rise Time all added significant unique variance (19.8% to 11.3%) when added at Step 3, a further model was created to further examine the influence of each AT variable. Frequency (with the greatest amount of unique variance) was added at Step 3, followed by Duration and then Rise Time.

Table 6-18 Results of Regressions exploring the unique variance in Word Score accounted for by Age (months), NVIQ and AT measures

Word Score	b	SEb	β	*t*	*p*	ΔR^2	*p*
Model 5							
Age	.041	.170	.034	.241	.811	.101	.063
NVIQ	1.459	1.403	.159	1.040	.307	.157	.014
Frequency	-11.899	7.787	-.376	-1.528	.137	.198	.002
Duration	-.106	.057	-.273	-1.867	.072	.057	.071
Rise Time	-.018	.050	-.076	-.353	.726	.002	.726

The pattern of results was similar across both Word and Syllable level analyses. The predictor of Age just missed significance throughout the models *(p* range = .055 - .083*).* The second entered variable of NVIQ contributed significant amounts of unique variance (ΔR^2 range = 15.5% - 16.7%). Rise Time contributed further unique variance (Word: 7.7%; Syllable 11.3%) which was a significant change at Syllable Level (*p* = .021). Duration also contributed significant unique variance when entered at Step 3 (Word: 12.2%, *p* = .016; Syllable: 12.4%, *p* = .015) as did Frequency (Word: 13.4%, *p*= .013; Syllable: 19.8%, *p* = .002). Intensity did not contribute significantly to either analysis. When the co-varying AT tasks were added sequentially to the model (Model 5), the additional contribution of Duration (Word: 6.7%, Syllable 5.7%) was found to just miss significance (*p* = .067, .057), whilst Rise Time failed to add significant extra variance when entered last into the model.

6.15 Effect of Entrainment on Nonword Repetiton

The AMC and DLD groups completed both an Unentrained and an Entrained version of the task. We hypothesised that listening to an entraining beat prior to the stimulus would facilitate more accurate repetition.

6.15.1 Effect of Entrainment on Word- and Syllable-Level Accuracy

In order to investigate whether listening to an entraining beat impacted on accuracy of Nonword Repetition at Word and Syllable Levels, we conducted a series of repeated-measures ANOVAs (2 x 2 – Group[AMC, DLD] x Entrainment [Unentrained, Entrained] for DVs of Word Score, then Syllable Score.

Table 6-19 Word and Syllable Scores by Entrainment for AMC and DLD groups

Score	**AMC Mean (SD)**		**DLD Mean (SD)**	
	Unentrained	**Entrained**	**Unentrained**	**Entrained**
Word Score	22.29 (5.552)	22.88 (5.551)	8.43 (4.653)	10.36 (4.814)
Syllable Score	81.134 (11.324)	82.494 (9.515)	49.256 (11.852)	52.579 (11.558)

There was no main effect of Entrainment at either Word (*F*(1, 36) = 3.887, *p* = .056) or Syllable Level (*F*(1, 36) = 3.666, *p* = .064) although the trend was towards the Entrained scores being higher at both levels – see Figure 6-16.

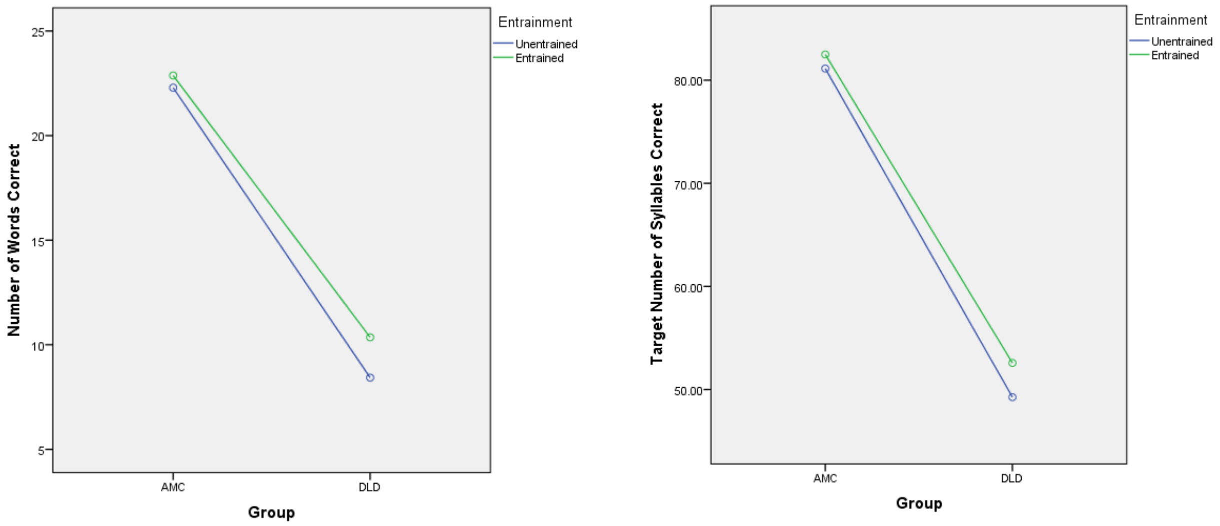


Figure 6-16 Graphs displaying Entrainment by Group for Words and Syllables Correct

When the ANOVAs were computed individually by group, there was no significant effect of entrainment at Word or Syllable Level for either the AMC group (Word: *F*(1, 23) = .676, *p* = .419; Syllable: *F*(1, 23) = 1.304, *p* = .265, or the DLD group (Word: *F*(1, 13) = 2.833, *p* = .116; Syllable: *F*(1, 13) = 1.792, *p* = .204.

Entrainment did not therefore result in significantly more accurate Nonword Repetition at either Word or Syllable levels of accuracy.

6.15.2 Interaction of Entrainment with Wordlength

To discover whether the effect of entrainment differed by Wordlength, we conducted a Repeated measures ANOVA (2 x 3 x 2 – Group [AMC, DLD] x Wordlength [3, 4, 5] x Entrainment [Unentrained, Entrained]. The interaction between Wordlength and Entrainment was not significant, *F*(2, 72) = 1.811, *p* = .171, however there was a significant Entrainment*Wordlength*Group interaction, *F*(2) = 3.392, *p* = .039).

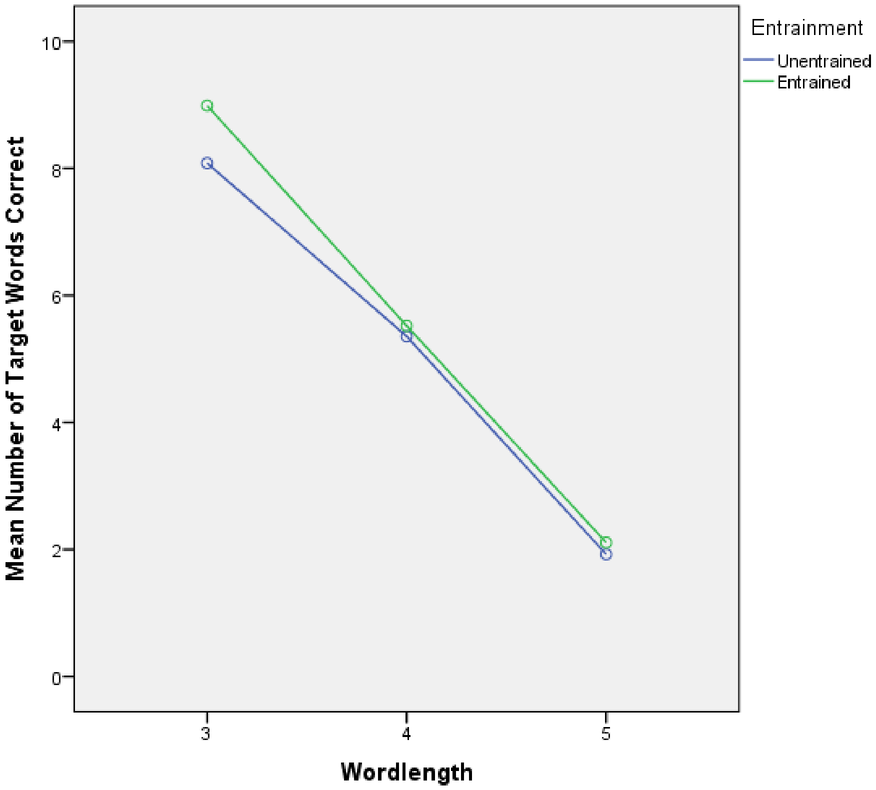
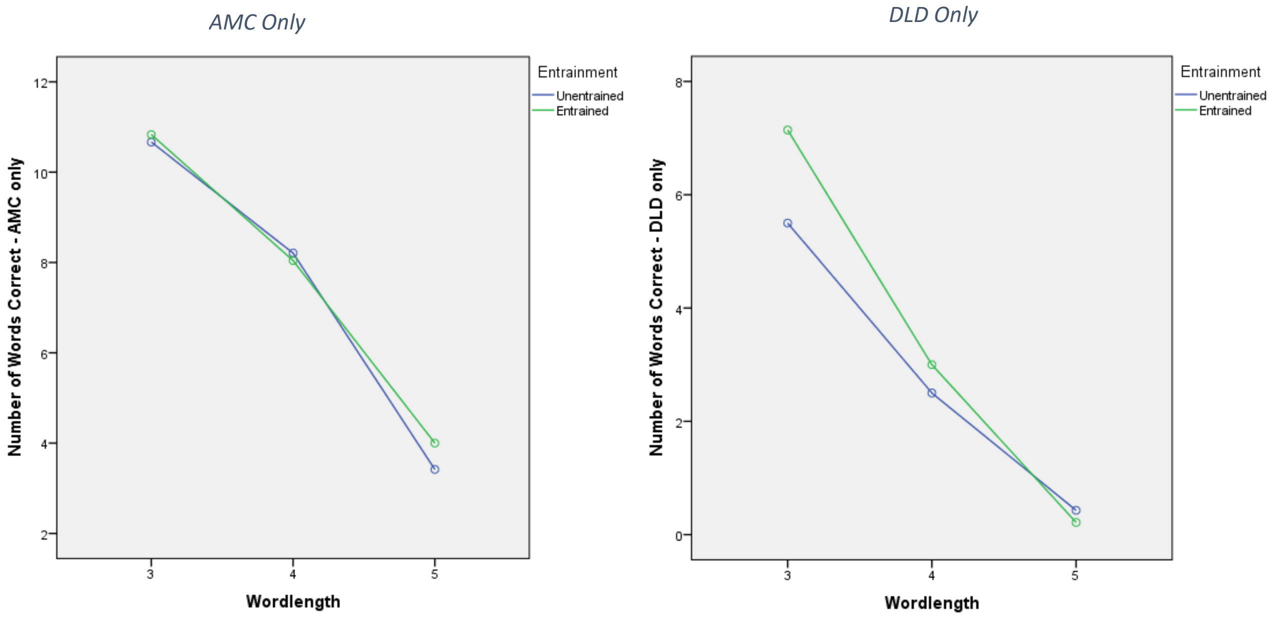


Figure 6-17 Graph showing Mean Number of Words Correct at each Wordlength by Entrainment condition

The Entrainment*Wordlength*Group interaction was investigated by running the ANOVAs again for each group individually. Entrainment * Wordlength was not significant for the AMC group, *F* (2, 46) = .975, p = .385, however it was significant for the DLD group *F*(2, 26) = 3.572, *p* = .043.

From the resulting graphs (Figure 6-18), we can see that Entrainment had a positive, but diminishing, effect for the DLD group for Wordlengths 3 and 4, whilst the scores at Wordlength 5 were similar in both versions of the task. In contrast, Unentrained and Entrained scores did not significantly differ for the AMC group.


AMC Only

DLD Only

Figure 6-18 Graphs showing Mean Number of Words correct at each Wordlength by Entrainment for AMC and DLD groups

Listening to an entraining beat did not therefore significantly affect repetition accuracy for the AMC group, and this did not differ by Wordlength. In contrast, Entrainment had a positive, but decreasing, impact on accuracy for the DLD group at Wordlengths 3 and 4, however was no longer beneficial at the longer Wordlength 5.

6.15.3 Effect of Entrainment by Stress Status

In order to investigate whether entrainment had a different effect on Stressed and Unstressed syllables, a repeated-measures ANOVA (2 x 2 x 2 – Group [AMC, DLD] x StressLevel [Stressed, Unstressed] x Entrainment [Unentrained, Entrained] – DV – Syllable score) was carried out.

There was no overall effect of entrainment, *F(*1, 36) = 2.012, *p* = .165, however there was a significant interaction between Entrainment and Stress Level *F*(1, 36) = 5.234, *p* = .028.

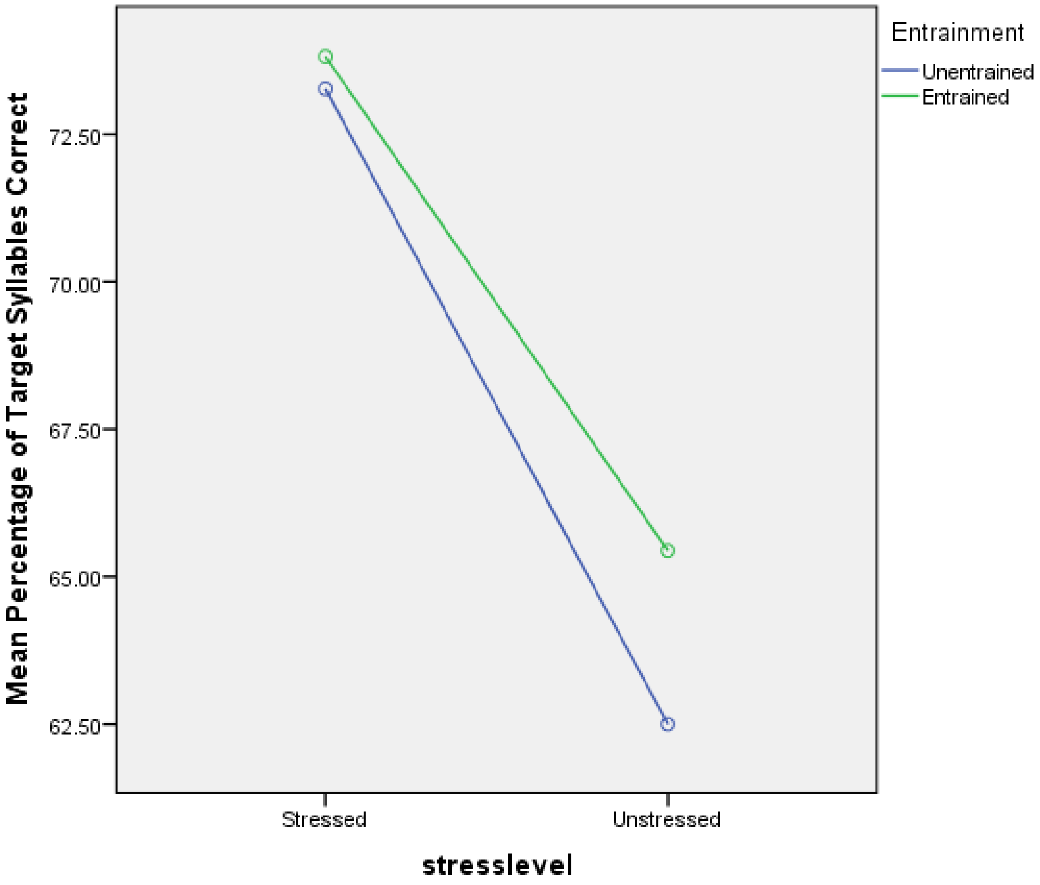


Figure 6-19 Graph showing Mean Percentage of Stressed and Unstressed Syllables correct by Entrainment

From Figure 6-19, we can see that this interaction arose because the difference between Entrainment levels was higher for the Unstressed than the Stressed syllables. When re-run for each group individually, there was a significant Entrainment*Stress interaction for the AMC group *F*(1, 23) = 4.404, *p* = .047, with accuracy rates for unstressed syllables being higher in the Entrained condition. In contrast, there was no interaction between entrainment and Stress Level for the DLD group *F*(1, 13) = 1.683, *p* = .217.

Table 6-20 Results of paired sample t-tests for percentage of Stressed and Unstressed syllables correct

	Unentrained	**Entrained**	***df***	***t***	***p***
Stressed Syllables					
AMC	87.612 (11.438)	86.921 (8.332)	23	.432	.669
DLD	58.929 (11.436)	60.714 (12.793)	13	-.740	.472
Unstressed Syllables					
AMC	78.974 (11.515)	81.019 (10.213)	13	-1.460	.168
DLD	46.032 (12.205)	49.868 (11.421)	23	-1.694	.104

Despite the interaction effect, a series of Paired Sample t-tests did not find a significant difference between Unentrained and Entrained Syllables whether Stressed or Unstressed for either participant group.

There was therefore a small beneficial effect of Entrainment for Unstressed syllables, however this effect was not sufficient to create a significant difference between Entrainment categories when examined individually.

6.15.4 Effect of Entrainment on Relative Stress Position

We then wondered whether the effect of Entrainment varied by Syllable Position relative to stress. We therefore conducted a repeated-measures ANOVAs (2 x 7 x 2 – Group [AMC, DLD] x Syllable Position [-3, -2, -1, STRESS, +1, +2, +3] x Entrainment [Unentrained, Entrained] – DV – Syllable Score).

Figure 6-20 Graphs showing Mean Percentage of Syllables Correct for each Syllable position relative to stress by Entrainment status

This revealed a significant effect of entrainment, *F*(1, 36) = 5.267, *p* = .028, with Entrained scores more accurate than Unentrained. There was also a significant interaction between Syllable Position, Entrainment and Group *F(*6) = 2.480, *p* = .024.

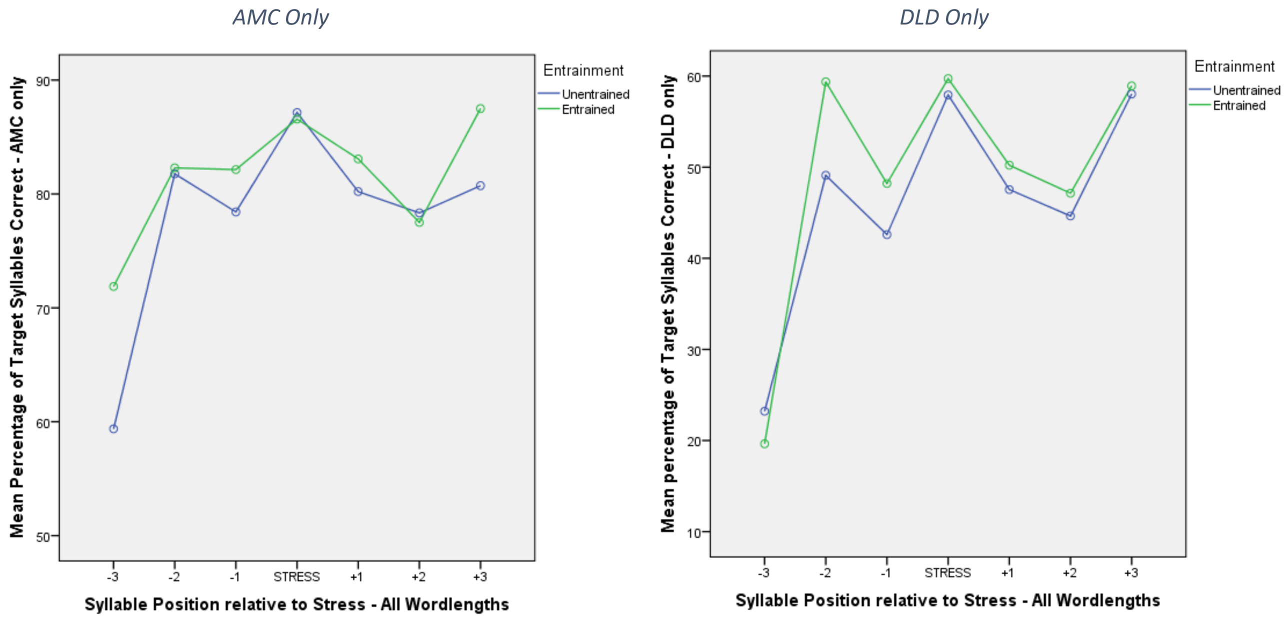


Figure 6-21 Graph showing Percentage of Target Syllables Correct for each Syllable Position Relative to Stress across Wordlengths for AMC and DLD groups

To investigate this interaction further, the ANOVAs were repeated for each group individually. For the AMC group, there was a significant main effect of Entrainment, *F*(1, 23) = 7.748, *p* = .011, but not for the DLD group, *F*(1, 13) = .940, *p* = .350. For neither group did Entrainment interact significantly with Syllable Position. Despite this, we can see a broad pattern emerging from the AMC graphs, suggesting that whilst Entrainment had no effect for stressed syllables, it was beneficial for the surrounding unstressed syllables (-1, +1) and also for the word-initial and word-final syllables (-3, +3).

6.15.5 Summary of Effect of Entrainment

We had hypothesised that listening to an entraining beat would provide a prosodic scaffold for the upcoming stimuli, resulting in more accurate repetitions. This was not what we found, however, as despite a trend towards slightly higher Entrained scores, there were no significant differences between accuracy rates in the Unentrained and Entrained versions at either word or syllable level. We did find evidence of some subtle influences of entrainment however. For the DLD group, the effect of entrainment differed across Wordlengths, so that there was a beneficial effect of Entrainment at shorter Wordlengths, which decreased as the words got longer, so that there was no longer a benefit at Wordlength 5. For the AMC group, however, the effect of Entrainment did not differ by Wordlength, but by stress status. Entrainment did not affect their repetition of Stressed syllables, however there was a significant beneficial impact on accuracy of Unstressed syllable repetition. From the Relative position to Stress analyses, this was predominantly driven by improved accuracy for syllables surrounding the stress (-1, +1) and for those syllables occurring word-initially and word-finally (-3, +3).

6.16 Summary of Findings from Nonword Repetition task

This task set out to systematically investigate the influence of the prosodic structure of the target on repetition of nonwords. We hypothesised that the presence and location of the stressed syllable within a nonword would influence the accuracy of repetition.

We found that, consistent with previous literature, children with DLD were poorer at repetition than their age-matched peers, and tended to be less accurate than language-matched younger children (although this was only significant at Wordlength 4). We also replicated an effect of Wordlength, whereby the longer the nonword, the less accurate its repetition, and this was true for all three groups.

By systematically varying the location of the stressed syllable within each Wordlength, we were able to investigate whether the different resulting prosodic structures (e.g. Sww, wSw, wwS) resulted in different accuracy levels. We found that location of the stressed syllable did affect repetition accuracy but that this differed by Wordlength and participant group. Stress location was significant for the DLD and YLC groups at Wordlength 3, for the YLC group only at Wordlength 4, and for the AMC group only at Wordlength 5.

We also found that the position of each syllable within a word (without taking account of stress location patterns) also impacted on how accurate repetition of that syllable was. This was significant for the AMC group at Wordlength 3, for the DLD and YLC groups at Wordlength 4, and for all groups at Wordlength 5. We saw that as the words increased in length, there developed an increasing advantage for syllables occurring word-initially and word-finally when compared against the accuracy of mid-word syllables.

We then investigated how the location of the stressed syllable interacted with syllable location in affecting accuracy of syllable repetition. We discovered that the stressed syllable within each prosodic structure tended to be the more accurately repeated, but that with increases in length, the effect of syllable location superceded the preserving factor of stress such that at Wordlength 5, the presence of the stressed syllable at position 5:3 was insufficient to overcome the negative impact of being a mid-word syllable.

Stress had a highly significant impact on syllable repetition accuracy for all groups, with stressed syllables more accurate than unstressed syllables. When examined in detail, this varied by participant group and Wordlength, with stress a significant factor the DLD group at Wordlength 3, AMC and YLC groups at Wordlength 4 and for all three groups at Wordlength 5. We also discovered that the effect of Wordlength was different for Unstressed and Stressed syllables. Unstressed syllables decreased in accuracy as Wordlength increased, mirroring overall accuracy levels. In contrast, Stressed syllable accuracy remained relatively constant across Wordlengths for AMC and YLC groups (although there was a small decrease in accuracy for AMC from Wordlength 3 to Wordlengths 4 and 5). The DLD group showed a different pattern, however, with a sharp decrease in Stressed syllable accuracy from Wordlength 3 to Wordlength 4, but then with accuracy rates rising at Wordlength 5 to be similar to those of Wordlength 3. Stress status therefore interacts significantly with Wordlength for this group.

We also examined the effect of proximity to the stressed syllable on syllable repetition accuracy. We found that pre-stress (-1) syllables tended to be less accurate than stressed syllables at Wordlengths 3 and 4. This was not the case at Wordlength 5, however, where proximity to the stressed syllable (-1 and +1) had a preserving effect. As observed previously, we could also see an interaction between position relative to stress and syllable location within words, as syllables falling far from stress, but in word-initial or word-final position (e.g. +3 at Wordlengths 4 and 5) were also relatively accurate compared to other syllables. Position relative to stress also varied according to participant group and Wordlength with the effect significant for the DLD group at Wordlength 3, the YLC group at Wordlength 4 and for all groups at Wordlength 5.

We subsequently investigated whether accuracy in Nonword Repetition was related to measures of sensitivity to the acoustic parameters of Rise Time, Duration, Frequency and Intensity. We found that greater sensitivity to changes in Rise Time, Duration and Frequency correlated with more accurate Nonword Repetition. Acoustic sensitivity to Rise Time, Duration and Frequency were all found to contribute significant unique variance to Nonword Repetition scores once Age and NVIQ had been accounted for, with Frequency making the greatest contribution.

Finally, we discovered that listening to an entraining beat did not significantly affect accuracy of Nonword Repetition overall, although there was evidence of some subtle beneficial effects. For the DLD group, entrainment facilitated accuracy at shorter Wordlengths, whilst for the AMC group, a beneficial effect on Unstressed syllable accuracy was found for syllables surrounding stress (-1, +1) and for word-initial and word-final syllables (-3, +3).

Our results therefore demonstrate that the prosodic structure of words has a complex and direct impact on repetition with accuracy rates at both word and syllable level influenced by the position of the stressed syllable and overall stress status. The influence of prosodic structure also varied according to Wordlength, suggesting an interaction between prosody and memory systems. Influence of prosodic structure by Wordlength also varied according to participant group, suggesting that the interaction of prosody and memory varies according to developmental characteristics.

6.17 Discussion

This task set out to discover whether stress patterns influenced children's repetition accuracy for nonwords, systematically examining the influence of stress and the location of stress at three- four- and five-syllable wordlengths. Analysis at word- and syllable-level points to a systematic influence of stress upon the accuracy of children's responses in NWR tasks across wordlengths and linguistic units.

Consistent with previous literature, the DLD group were less accurate in their repetitions than the AMC group (e.g. Dollaghan & Campbell, 1998) and the YLC group (Gathercole & Baddeley, 1990). Our participants with DLD therefore demonstrated the predicted difficulties with NWR tasks.

There were significant effects of wordlength on accuracy for all groups – the longer the target nonword, the lower the accuracy level. Again, this is a previously well-attested effect (Dollaghan & Campbell, 1998; Gathercole et al., 1991). The children with DLD made errors at Wordlength 3, whilst AMC children began to show errors at Wordlength 4, which is similar to the group-wordlength error relationship found in other studies (Marton & Schwartz, 2003). The wordlength effect in NWR has often been taken to indicate the primary influence of phonological memory on performance in NWR tasks, with longer wordlengths representing a greater memory load and thus a greater likelihood that storage capacity will be exceeded.

Further evidence for the role of PSTM in NWR came from our analysis of Syllable Location. The influence of Syllable Location on accurate repetition varied between Wordlengths and Group. There was no overall effect of Syllable Location at Wordlength 3, whilst there was a significant effect for DLD and YLC groups at Wordlength 4, and for all groups at Wordlength 5. Inspecting the response curves for the groups indicates that at longer Wordlengths, children were beginning to show primacy and recency effects in their response, with Syllables 1 and 4/5 having better accuracy figures than intervening syllables. Here, as the memory load increased, primacy (AMC – Wordlength 5) and recency effects (DLD, YLC – Wordlength 4, All groups – Wordlength 5) became more apparent, suggesting that the task demands were beginning to exceed capacity. Our results demonstrated that primacy and recency effects can be observed within words as well as across word lists.

There were also significant effects of stress in our results. Stressed syllables were more accurately repeated than unstressed for all groups, and at most wordlengths (except AMC Wordlength 3, DLD Wordlength 4). This suggests that stressed syllables have a privileged status in phonological memory, resulting in more accurate recall. Furthermore the rate of stressed syllable accuracy was relatively constant across wordlengths for AMC (although slight decrease Wordlength 3-4) and YLC groups, whilst the unstressed syllable rate declined sharply for all groups with each wordlength increase. This means that the extent of privileging of the stressed syllable increases the greater the PSTM load, suggesting that at longer wordlengths, the stressed syllable may be preserved at the expense of unstressed syllables.

Further evidence for the influence of stress arises from our analysis of syllable accuracy in relation to proximity to the stressed syllable. Pre-stress (-1) syllables tended to be less accurate than stressed syllables at Wordlengths 3 and 4, which is consistent with previous observations (e.g. Roy & Chiat, 2004). Previous studies with short 2/3-syllable words have found that post-stress (+1) syllables were better repeated than -1, however this was not the general pattern of our data (exception – Wordlength 4 YLC only). Instead, we found that as wordlength and memory load increased (i.e. at Wordlength 5) the stress and its surrounding syllables (+1, -1) were the most accurate. This suggests that as memory capacity is exceeded, the stressed syllable retains its privileged position in PSTM, and that it provides a scaffolding structure for the retention of the surrounding syllables. This provides a parallel source of accuracy to that provided by the recency effects to create the patterning at Wordlength 5 whereby -1, STRESS, +1 and +3 tended to be the more accurate syllables.

The primary retention of stressed syllables within the phonological loop as found in our data, may be due to their enhanced acoustic signature (louder, longer) acting as an anchor for phonological analysis, giving these elements priority within the phonological loop. Gathercole & Baddeley (1990) suggested that if the acoustic and segmental analysis of phonological material was ‘noisy’ (p.357), in children with DLD then resulting representations would be less discriminable at the point of retrieval. Since children with DLD have poorer sensitivity to acoustic dimensions (AT tasks), it can be speculated that this results in ‘noisy’ analysis. In such a state of noise, it seems likely that the more prominent (i.e. stressed) aspects of the signal will be preferentially preserved when the system is overloaded (e.g. when capacity is challenged).

Memory systems are not the only factor, however, as we also found an effect of rhythmic structure, even when wordlength was held constant. Rhythmic structure has previously been found to affect repetition accuracy (Chiat & Roy, 2007; Roy & Chiat, 2004), but this is the first study to systematically manipulate rhythmic structure at longer wordlengths for older children. The effect of rhythmic pattern differed by group. For the DLD group there was an effect at Wordlength 3; for YLC at Wordlengths 3 and 4, whilst for AMC Groups there was an effect at Wordlength 5 only. This set of results suggests a developmental interaction between rhythmic and memory systems, whereby when the target is comfortably within memory capacity (Wordlengths 3 and 4 for the AMC group) – stress syllable position has little effect with all rhythms produced equally well. Conversely, at Wordlength 5 (3 and 4 DLD, YLC) when memory capacity was more stretched, the influence of stress position could be seen in the wider range of scores of produced. As memory capacity was exceeded (Wordlength 4, DLD; Wordlength 5 DLD, YLC), stressed syllable position was no longer relevant as the memory load was too great.

The specific rhythmic patterns that were better/more poorly repeated were also of interest. We had hypothesised that internalised rhythmic templates would support task performance. If that were the case, we would expect the most frequently occurring patterns in English to be more accurately repeated. This was not what we found, however. 3:3 (wwS) is by far the least commonly occurring three-syllable pattern (7%), however it was the most accurately repeated pattern for DLD and YLC groups. Conversely, 4:2 (wSww) was the least accurate pattern for the YLC group, yet is the most frequently occurring four-syllable pattern (42.7%). This suggests that children are relying less on rhythmic templates for task completion than we had hypothesised. Accordingly, the impact of rhythmic characteristics may lie in the on-line processing of specific acoustic characteristics of the stimulus rather than in the redintegration process from long-term knowledge.

We predicted that acoustic sensitivity would impact on accuracy of repetition and Rise Time, Duration and Frequency were all found to contribute significant amounts of unique variance to Nonword Repetition Score above the influence of age and NVIQ. Of these co-varying factors, Frequency provided the greatest contribution, followed by Duration whilst Rise Time did not provide further contribution when entered last into the model. These results point towards acoustic sensitivity as a significant factor in repetition accuracy.

We had hypothesised that listening to an entraining rhythm would result in more accurate repetition. Despite the lack of an overall effect of Entrainment on accuracy, there were some subtle beneficial effects. Repetition was more accurate at shorter Wordlengths in the Entrained condition for the DLD group. This suggests that when the target is within PSTM capacity, then entrainment can support repetition for this group. This may be through provision of a skeleton rhythmic structure on which the specific phonology can be grafted (cf Morgan & Saffran 1995), or by creating a temporal expectancy for the arrival of the stimulus. The AMC group showed a pattern for Entrainment in which certain Unstressed syllables (-1, +1; word-initial and word-final -3, +3) benefitted in the Entrained condition. If increases were due to temporal expectancy, then we would have expected increased accuracy across all syllables, however the selective effect at phonological anchor points already noted to be privileged in PSTM is more suggestive of structural scaffolding.

Rhythmic patterning therefore has a significant influence on accuracy in NWR. Stressed syllables appear to occupy a privileged position in PSTM, resulting in more accurate repetition compared with unstressed syllables. At long wordlengths, when memory capacity is exceeded, the stressed syllable appears to have a protective effect on surrounding syllables whilst syllables outside its immediate proximity decrease in accuracy. Children with DLD already had poorer accuracy than AMC children at Wordlength 3, with increasing difficulties as words became longer, suggesting that they may become reliant on the stressed syllable preservation effect at shorter wordlengths than TD children, with concurrent underspecification of more vulnerable unstressed syllables. A practical implication of this is that they are likely to need support in specifying representations of the longer words which can characterise academic vocabulary at KeyStage 2 and beyond. The Entrainment data suggests that provision of a direct rhythmic scaffold could support more accurate acquisition of such words.

7 Experimental Task 3 – Lexical Stress

The next aspect of language development we chose to investigate was also at the lexical level. The Nonword Repetition task investigated how lexical stress patterns affected the processing of novel material. Nevertheless, as new vocabulary is acquired and represented permanently in the language system, the internal rhythmic structure of each word must also form part of that permanent record. We hypothesised that if children with DLD were less sensitive discriminators of rise time, and hence of stress, that they would find it difficult to incorporate a stable representation of the rhythmic stress pattern of a word into their mental lexicons. To explore this idea, we chose to include an experiment examining children's representations of lexical stress for items likely to be already established in their vocabulary.

7.1 Theoretical Basis of the Task

In language development, the characteristic pattern of strong and weak syllables assigned to individual words appears to be encoded along with other aspects of the word's structure from as early as 7 months (Curtin, Mintz, & Christiansen, 2005). Lexical and rhythmic knowledge therefore appear to be integrated from the earliest stages of language-learning.

As outlined in Section 1.2.3, we previously found that children with DLD had poorer representations of lexical stress than TD peers, and that their scores on our experimental task related to measures of rise time discrimination (Richards & Goswami, 2015).

To further explore the lexical stress representations of children with DLD, we decided to build on this earlier result for our third experimental task. Previously, we had used single words to investigate lexical stress. In natural speech, however, words usually occur as part of a sequence rather than in isolation. We therefore considered whether the global rhythmic context in which a word occurred could influence children's judgements of lexical stress.

There is some evidence that acoustic patterns in preceding material can influence listener responses to the final, target, stimulus. This is thought to be caused by listeners using the patterns present in the ongoing stimulus to make predictions about the nature of the upcoming material. Thomassen (1982) described this process of inducing expectations as "controlled anticipation" (p1597). His experiments used non-linguistic stimuli (e.g. high and low tone patterns) to investigate the perception of accent and demonstrated that the characteristics of the preceding tone sequence affected judgements of sequence regularity. This notion is parallel to that previously discussed under dynamic attending theory. If dynamic attending theory considers that expectancies are set up as to the ‘when’ of the next stimulus, then ‘controlled anticipation’ could be considered as manipulating the expectation of ‘what’.

Recent studies have demonstrated similar effects using prosodic manipulation. When adults are played word sequences that induce a strong Sw sequence, they are more likely to segment target ambiguous word sequences (e.g. note book worm) according to the preceding Sw pattern (i.e. as ‘notebook’ (Dilley & McAuley, 2008) suggesting that the global (or ‘distal’ in their terminology) rhythmic pattern can influence more local rhythmic decisions. Manipulating the fundamental frequency patterns of a carrier sentence towards Sw or wS can induce expectations of a Sw or wS pattern in the final target word (Brown, Salverda, Dilley, & Tanenhaus, 2015; Niebuhr, 2009), whilst judgements of which acoustically neutral final word is stressed are affected by the preceding rhythmic sequence (Martin, 1970). Preceding rhythmic patterns have also been shown to affect word segmentation of lexically ambiguous syllables (Brown, Salverda, Dilley, & Tanenhaus, 2011). These studies of adult listeners demonstrate that mature language-users appear to use rhythmic patterns derived from preceding stimuli in order to derive an expectation of what will arrive next. This expectation in turn influences what they report perceiving in the signal.

We hypothesised that a similar influence of rhythmic context might prevail on children’s perceptions of lexical stress. Infant-directed speech is markedly more rhythmic than adult-directed speech and other aspects of young children’s oral culture (e.g. nursery rhymes and rhyming books) are also highly rhythmic in nature. We speculated that this marked rhythmicity in early development could aid language development by providing a strong rhythmic context for individual words, thus drawing greater attention to their prosodic structure. We thought that if the prevailing rhythm coincided with the rhythmic structure of a target word, its representation would be strengthened due to the congruence of cues. In the context of an experimental task, we hypothesised that this rhythmic congruence might translate as more accurate judgements of lexical stress patterns. In contrast, we thought that if the target word did not conform to the prevailing rhythm, this may be disruptive to judgements of stress pattern.

7.2 Devising the Task

This task was designed to investigate whether the preceding rhythmic context of a sentence influenced the ability of children to determine whether the *final* word of that sentence was correctly stressed.

7.2.1 Creating the Carrier Sentences

There were two major considerations in constructing the carrier sentence: the rhythmic construction; and grammatical and semantic consistency between sentences.

We wanted minimal differences in grammar and semantics between the two different rhythms and so we used as similar a construction as possible to create two sets of rhythmic expectations:

1. Jack is reading books about a …… (*completed by target*)
2. The boy is reading books about …… (*completed by target*)

Both sentences have an equivalent syntactic structure with the only difference being in the subject noun phrase (Jack; the boy). The different noun phrases enable the induction of a different rhythmic expectation for the final word, whilst keeping the number of metrical feet per introduction constant. The 'Jack' sentence creates a carrier of four Sw trochaic metrical feet prior to the target word. The 'boy' sentence creates a carrier of four wS iambic metrical feet prior to the target word.

Table 7-1 Rhythmic structure of the carrier sentences for Lexical Stress task, using example target 'chicken'

Target Stress	**Rhythm**	**Stress pattern (sentence)**										
		w	**S**	**w**	**S**	**w**	**S**	**w**	**S**			
CorrStress	Trochaic		JACK	is	REA	ding	BOOKS	a	BOUT	a	**CHI**	cken
CorrStress	Iambic	The	BOY	is	REA	ding	BOOKS	a	BOUT	**CHI**	ckens	
IncorrStress	Iambic	The	BOY	is	REA	ding	BOOKS	a	BOUT	chi	**CKENS**	
IncorrStress	Trochaic		JACK	is	REA	ding	BOOKS	a	BOUT	a	chi	**CKEN**

Note: grey shaded cells indicate the location of the target word; capitals indicate stressed syllables

The first manipulated variable was the location of the primary stress of the final word. Each sentence would be completed with a target noun which either carried the correct stress (CorrStress) pattern (Rows 1 and 2; primary stress on first syllable) or an incorrect stress (InCorrStress) pattern (Rows 3 and 4; primary stress on second syllable). This allowed us to investigate the question of whether children were able to determine which was the correct stress pattern for the target word.

The second manipulated variable was the location of the primary stress with regard to the preceding rhythmic expectation set up by the carrier sentence. As can be seen in Table 7-1, the location of the stressed syllable of the final target word either continued the alternating pattern (Rows 1 and 3) or disrupted the preceding pattern (Rows 2 and 4). This allowed us to investigate whether the rhythmic expectation of the preceding sentence affected how accurately children were able to determine the correct stress pattern.

The combination of these variables created four different conditions:

Table 7-2 The structure of the four Lexical Stress conditions with their definitions

Conditions		**Abbreviation**	**Stress pattern (sentence)**										
Stress	**Rhythm**		**w**	**S**	**w**	**S**	**w**	**S**	**w**	**S**			
Correct	Congruent	**CorrStress-Con**		JACK	is	REA	ding	BOOKS	a	BOUT	a	**CHI**	cken
Correct	Incongruent	**CorrStress-InCon**	The	BOY	is	REA	ding	BOOKS	a	BOUT	**CHI**	ckens	
Incorrect	Congruent	**InCorrStress-Con**	The	BOY	is	REA	ding	BOOKS	a	BOUT	chi	**CKENS**	
Incorrect	Incongruent	**InCorrStress-InCon**		JACK	is	REA	ding	BOOKS	a	BOUT	a	chi	**CKEN**

7.2.2 Choosing the target words

In choosing the target words, we considered aspects of the stress and phonological structure of the words as well as attempting to control for variables such as familiarity and age of acquisition, whilst ensuring that the words were likely to be known by our participants.

In order to investigate these aspects systematically, we used the on-line MRC Psycholinguistic database. The database was searched for words matching the following criteria: 1. words of two syllables; 2. nouns; 3. carrying primary stress on the first syllable. This resulted in a list of 543 words.

English stress rules mean that unstressed vowels are frequently realised as [ə]. If such a syllable were stressed (as required by the paradigm) it would necessitate a change in vowel quality, providing an additional, phonemic cue to stress status. In order to avoid reliance on this phonemic cue, all words whose unstressed syllable contained [ə] were removed. There are also several pairs of words in which the noun has a Sw pattern and a corresponding verb has a wS pattern (for example ‘an OBject’ (Sw noun) and ‘I obJECT’ (wS verb). Deciding which of these was correct within our paradigm would involve additional linguistic processing as to whether a noun or verb was intended, and so these words were also removed from the list. This left 231 words.

From the remaining words, note was taken of ratings of Age of Acquisition, Familiarity, Imageability and Concreteness in order to find words occurring within the top 50 in each of those categories. This was in order to ensure maximum familiarity with the target words. There were 18 words which occurred in the top 50 across all the categories. The frequency value of these 18 words was then calculated based on the British National Corpus figures (Leech, Rayson & Wilson, 2001). Six outliers were identified and removed leaving 12 words in the list. These were the words used in the final task (2 as practice items and 10 as targets).

We therefore had a list of target words which we could be confident would be familiar to the children, had a grammatical status which would fit with the target sentence and whose phonological structure would not be altered by a change in stress structure since both syllables contained vowels capable of carrying primary stress.

7.2.3 Creating the Stimuli

During recording of the stimuli, a regular beat was induced in the speaker using a priming metronome beat in one ear (inaudible on the recording) set to 120 beats per minute (equivalent to 500ms between beats). The stimulus was then spoken so as to align the stressed syllables of the carrier phrase with the beats at 500ms inter-stress intervals[21]. The timing of the target word was consistent with the overall rhythmic structure of 250ms between syllables. This meant that for the Incongruent sentences, the syllable rate remained constant, whilst the stressed syllable rate was disrupted. For CorrStress-InCon sentences, this resulted in a final stressed syllable interval of 250ms, whilst for InCorrStress-InCon sentences, the final stressed syllable interval was 750ms (See Figures 7-1 and 7-2). The precision of this timing was then verified and adjusted as necessary using Audacity software.

---

[21] The stressed syllable rate of 500ms here is different from others in the task battery because of the nature of the rhythm used. Other tasks had several syllables between beats, meaning that a 500ms interval was too rapid for accurate manipulation. Conversely this task had only two syllables per foot. When this reduced foot pattern was recorded at 750ms intervals, the resulting recording sounded excessively slow and unnatural. As a result, we chose to keep the syllable rate of 250ms intervals constant between the tasks, resulting in a 500ms stressed syllable interval for this task.

For the Entrained version of the task, the entraining segment had a regular rhythm, following either a Sw alternation ('Jack' sentences) or wS alternation ('the boy' sentences). The number of pulses in the entraining portion matched the number of syllables in the target sentence.

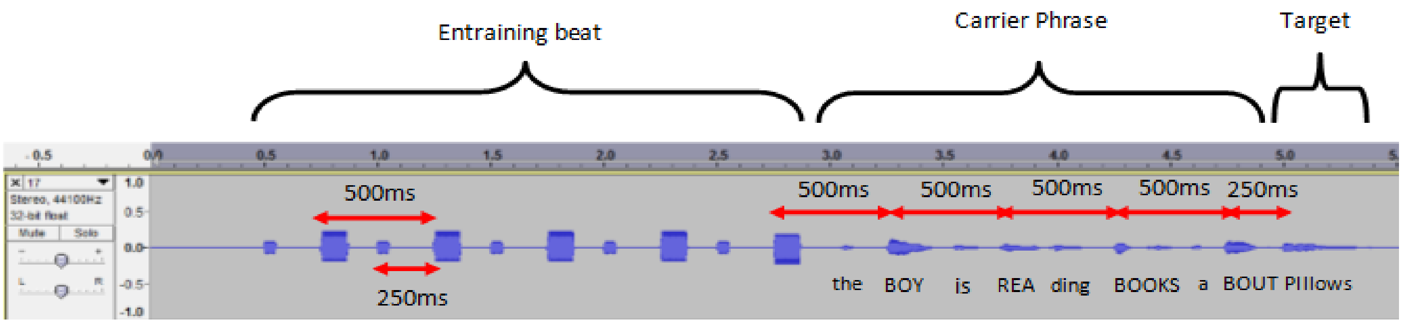


Figure 7-1 Soundwave of CorrStress-Incongruent sentence 'the boy is reading books about pillows'

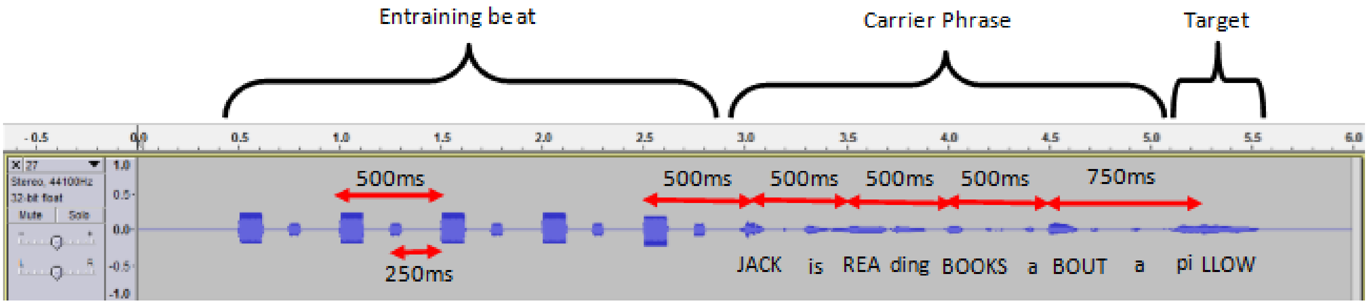


Figure 7-2 Soundwave of InCorrStress-Incongruent Sentence 'Jack is reading books about a pillow'

Each token was recorded and digitised individually. With ten target words across four stimulus types this resulted in forty individual recordings. These recordings were then split into two lists of twenty items each. Each word occurred twice in each list, once in each stress pattern, with the order of occurrence randomised. The stimulus type was balanced across the lists so that there were five CorrStress-Con, five CorrStress-InCon, five InCorrStress-Con and five InCorrStress-InCon examples in each list. The order of stimulus types within each list was also randomised. An Entrained and Unentrained version of each list was created.

7.2.4 Procedure

The task was contextualised by talking about stress as the 'beat pattern' in words and about how that could sound right or wrong, for example 'TAble' sounded right, whereas 'taBLE' sounded wrong. The children were told they were going to hear some sentences and they needed to decide whether the final word sounded right or wrong, so that 'Jack is reading books about a TAble' would sound right, and 'Jack is reading books about a taBLE' would sound wrong. If they thought the word sounded right, they should press the tick, and if they thought the word sounded wrong, they should press the cross.

Whilst each recording played back via the laptop, a picture of the relevant target word (e.g. a blanket) was displayed on the laptop screen.

Each presentation of the experimental task was preceded by four practice trials, during which feedback was given. For entrainment trials, the children were told that they would hear some beeps before the words which were there to help them hear the pattern of the words.

7.3 Results

A child's score was calculated as the number of correct responses accurately identifying a target word as being correctly or incorrectly stressed. The maximum score for each list was 20, with a maximum score of 5 for each Condition[22].

7.3.1 Accuracy (Scores)

The first set of analyses focused on the accuracy of children's responses.

---

[22] One YLC child's scores were only partially recorded by the software and so he was removed from the analysis. An administration error resulted in one AMC child and one DLD child repeating the same list (e.g. listened to List 1 in both Entrained and Unentrained tasks). Since all other participants listened to both List 1 and List 2, their scores were therefore not directly comparable and so they were removed from the analysis.

7.3.1.1 *Effect of Group on Accuracy (Score)*

Based on previous results, we predicted that the DLD group would score less well than the AMC group, showing them to be less successful at identifying whether target stress patterns were correct.

We calculated the number of items correctly identified for each group (Score) and performed a one-way ANOVA (DV-Score) to investigate if there was a group difference.

Table 7-3 Results of one-way ANOVA by group -DV Overall score

Score **(max 20)**	**AMC** **Mean (SD)**	**DLD** **Mean (SD)**	**YLC** **Mean (SD)**	***df***	***F***	***p***
Score[a]	16.35 (3.113)	14.62 (3.262)	13.15 (3.66)	2, 53	4.851	.012

Note: a) AMC > YLC

There was a significant effect of group (*p* = .012) with Bonferroni pairwise comparisons finding the AMC group more accurate than the YLC group (*p* = .009). The Score of the DLD group fell between those of AMC and YLC and was not significantly different from either group (*p* = .432, .682 respectively).

The DLD group were therefore not significantly poorer at identifying whether target stress patterns were correct in this task compared with the AMC group.

7.3.1.2 *Effect of Condition on Score*

There were four Conditions used in the task (CorrStress-Con, CorrStress-InCon, InCorrStress-Con, and InCorrStress-InCon). The mean scores for each Condition are given in Table 7-4 and plotted graphically in Figure 7-3.

Table 7-4 Mean Scores for Condition by Group

Score **(max 5)**	**CorrStress-Con** **Mean (SD)**	**CorrStress-InCon** **Mean (SD)**	**InCorrStress-Con** **Mean (SD)**	**InCorrStress-InCon** **Mean (SD)**
AMC	4.82 (.501)	4.23 (1.02)	3.68 (1.52)	3.59 (1.59)
DLD	4.71 (.825)	4.07 (.997)	3.07 (1.385)	2.93 (1.542)
YLC	4.25 (1.07)	4.15 (.745)	2.6 (1.875)	2.15 (1.663)
All Groups	4.59 (.112)	4.15 (.126)	3.12 (.222)	2.89 (.219)

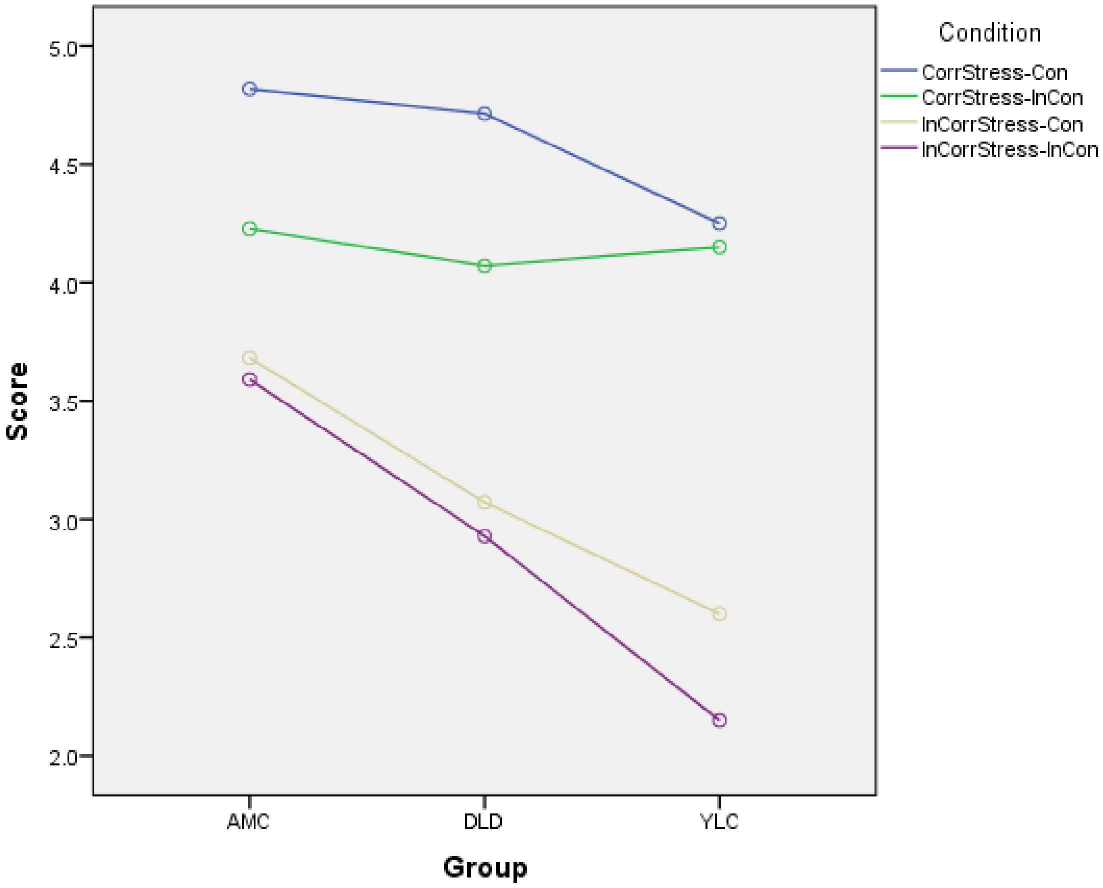


Figure 7-3 Graph showing mean scores by group for each Condition

We explored this data further by investigating our two specific predictions.

7.3.1.3 Effect of Congruence of Rhythm

Our stimuli varied according to whether the concluding word continued the rhythmic pattern established by the carrier sentence. We predicted that stimuli in which the resulting pattern remained Congruent (CorrStress-Con; InCorrStress-Con) would be more accurately judged than those in which the rhythm was disrupted (CorrStress-InCon; InCorrStress-InCon).

To test this hypothesis, we constructed two variables, by taking the mean of CorrStress-Con and InCorrStress-Con scores to create an overall Congruent score, and the mean of the CorrStress-InCon and InCorrStress-InCon scores to create an overall Incongruent score. These scores were then analysed using paired sample t-tests to quantify any overall difference between Congruent and Incongruent rhythmic phrases.

Table 7-5 Results of paired sample t-tests for Congruent and Incongruent rhythms by Group

Mean Score (Max 5)	**Congruent**	**Incongruent**	**df**	**t**	**p**
AMC[a]	4.25 (.869)	3.91 (.854)	21	2.417	.025
DLD[a]	3.893 (.836)	3.5 (.92)	13	2.242	.043
YLC	3.425 (1.067)	3.15 (.919)	19	1.565	.134
All Groups	3.8661 (.989)	3.536 (.938)	55	3.552	.001

Note: a) Con > Incon

Final words forming Congruent Rhythms were more accurately judged than Incongruent rhythms for AMC and DLD groups but not the YLC group, although the graphed results (Figure 7-4 below) show a trend in a similar direction.

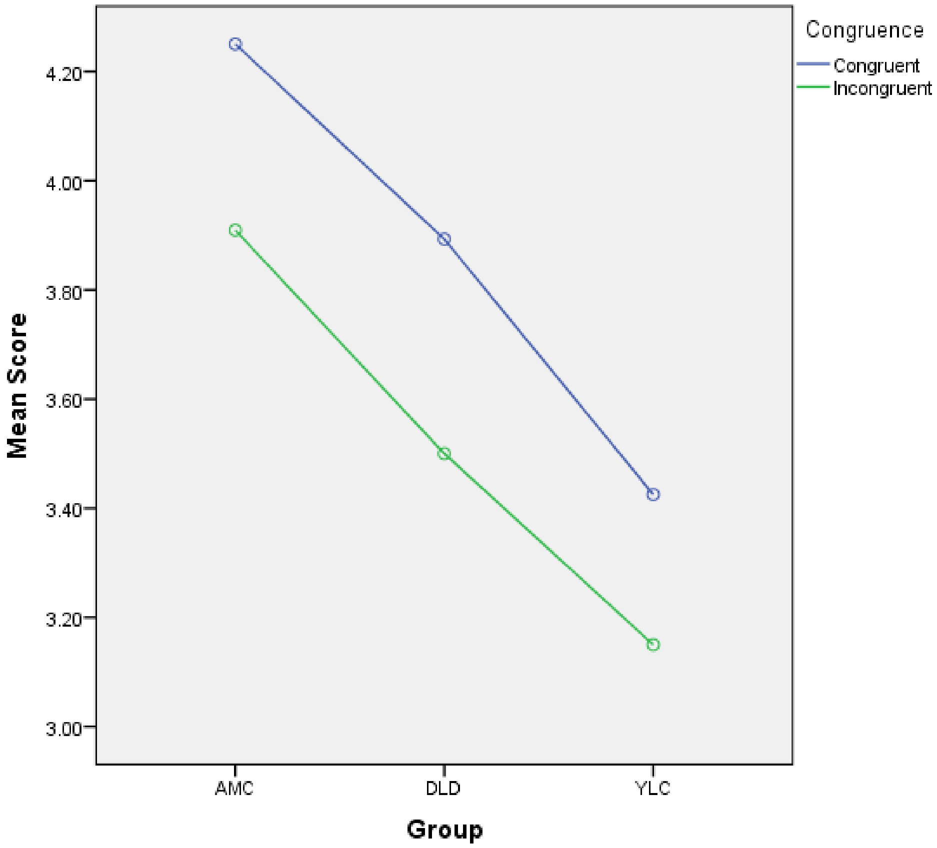


Figure 7-4 Graph showing the mean scores for Congruent and Incongruent concluding rhythms

Congruence of rhythm was therefore a significant factor in accuracy of response for the AMC and DLD groups, with stimuli in which the rhythm of the target word continued the pattern of the carrier phrase being judged more accurately than those in which the target word disrupted the carrier phrase rhythm. The difference in scores for the YLC group did not, however, reach significance.

7.3.1.4 *Effect of Target Stress (Correct or Incorrect Stress) on Score*

Inspecting Table 7-5 and Figure 7-5 (below) indicated a further potential influence on Score. Looking at Figure 7-5 in particular suggested that incorrectly stressed target words (beige and purple lines) were less accurately judged than correctly stressed words (blue and green lines). In other words, it seemed children were more likely to mistakenly accept incorrectly stressed words (e.g. judge chiCKEN as correct), rather than to reject correctly stressed ones (e.g. judge CHIcken as incorrect). To investigate this further, two variables were created: Correctly Stressed Target (Mean of CorrStress-Con and CorrStress-InCon scores) and Incorrectly Stressed Target (Mean of InCorrStress-Con and InCorrStress-InCon scores).

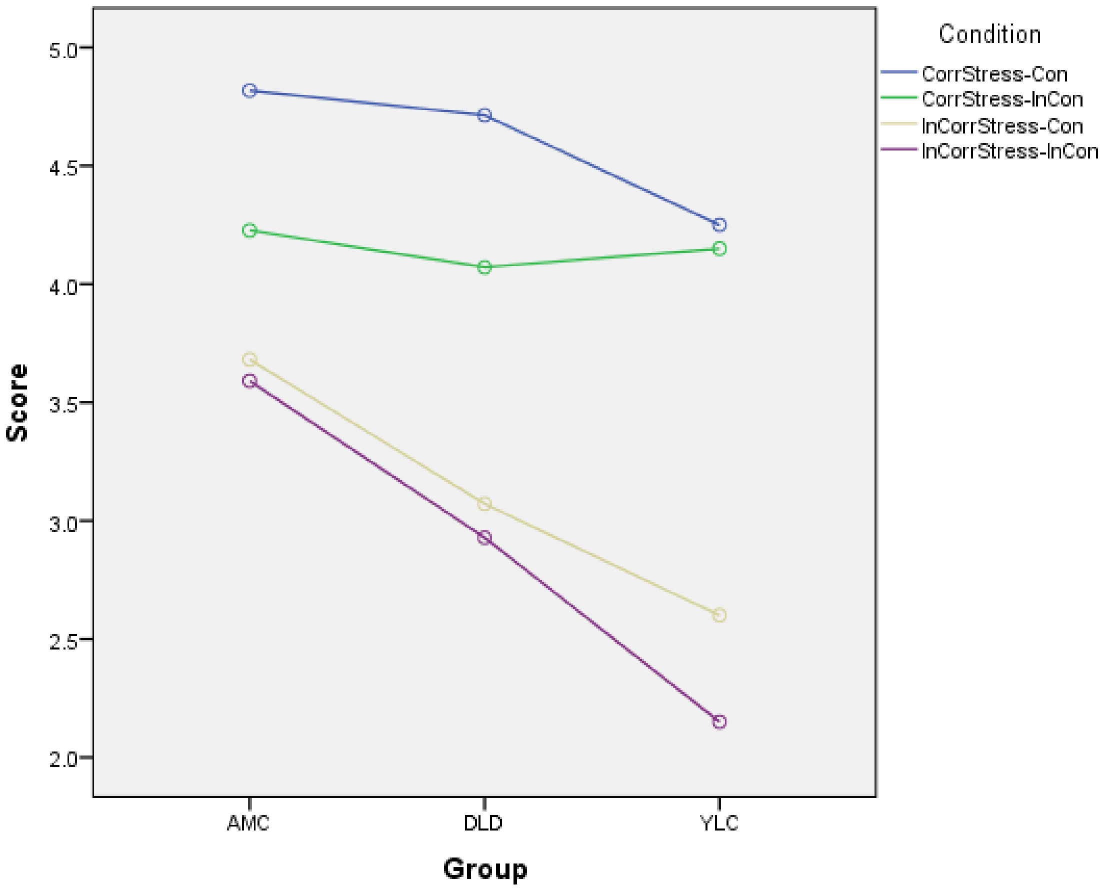


Figure 7-5 Graph showing mean scores by group for each Condition

Paired samples t-tests (Table 7-6) showed that Correctly Stressed Targets were judged significantly more accurately than Incorrectly Stressed targets by all groups.

Table 7-6 Results of paired sample t-tests for Correctly and Incorrectly Stressed Targets

Mean Score (Max 5)	**Correctly Stressed Target Mean (SD)**	**Incorrectly Stressed Target Mean (SD)**	***df***	***t***	***p***
AMC[a]	4.523 (.645)	3.636 (1.424)	21	2.709	.013
DLD[a]	4.393 (.813)	3.000 (1.373)	13	3.343	.005
YLC[a]	4.20 (.75)	2.375 (1.669)	19	4.460	.000
All Groups	4.375 (.728)	3.027 (1.577)	55	6.030	.000

Note: a) Correctly Stressed Target > Incorrectly Stressed Target

It was therefore the case that children were more accurate in their judgements when the target word was correctly stressed (e.g. CHIcken) than when it was incorrectly stressed (e.g. chiCKEN). From the graph of mean scores (Figure 7-6) it appeared that this difference was most marked for the YLC group. In order to quantify this, a measure of d' was calculated, using the number of correctly stressed words identified as correct as the hit rate, and the number of incorrectly stressed words wrongly identified as correct as the false alarm rate.

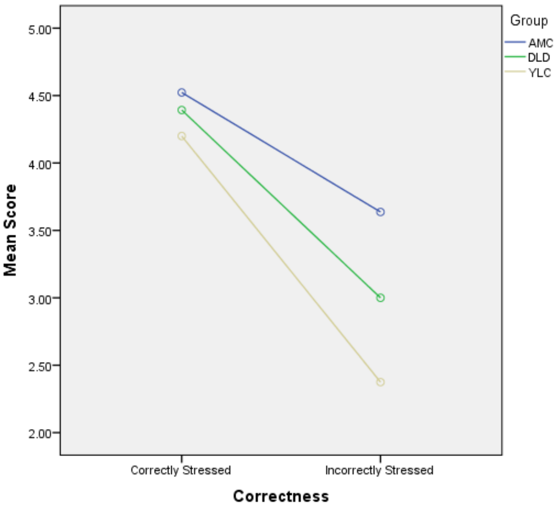


Figure 7-6 Graph showing Mean Scores for Correctly and Incorrectly Stressed Targets

A one-way ANOVA (dependent variable d') was conducted which showed a significant effect of group on d' ($F(2, 53) = 4.436$, $p = .017$) with the YLC group ($M = .975$) differing significantly from the AMC group ($M = 1.991$) ($p = .013$). The DLD group's d' measure ($M = 1.483$) did not differ significantly from either the AMC or the YLC group. This indicates that the YLC group were significantly less able to reject the incorrectly stressed words than the older AMC group, regardless of the rhythmic context in which the target word occurred.

Whether the target word was correctly stressed or not therefore had a significant influence on the accuracy of children's judgements for all three participant groups. Incorrectly stressed words (e.g. chiCKEN) were more likely to be accepted as correct, than correctly stressed words (e.g. CHIcken) were to be rejected. This effect was greatest for the YLC group.

7.3.2 Reaction Times

The Presentation software measured the time between the onset of the stimulus and the response button being pressed in units of .0001s. Because the different stimulus types resulted in different stimulus lengths (see Figures 7-1 and 7-2), it was necessary to adjust the time recorded by adding or deducting .2500s to some Conditions so that valid comparisons could be made. This effectively equalised the timing of the onset of the final stressed syllable. For Entrained versions, the entrainment portion was also deducted from the length.

7.3.2.1 Effect of Condition on Reaction Times

Mean Reaction Times were calculated for each group for each Condition – see Table 7-7.

Table 7-7 Mean RTs for each group by Condition

Reaction Times	**CorrStress-Con Mean (SD)**	**CorrStress-InCon Mean (SD)**	**InCorrStress-Con Mean (SD)**	**InCorrStress-InCon Mean (SD)**
AMC	39708 (6560)	43269 (8566)	43136 (9954)	40685 (7204)
DLD	44998 (14124)	44130 (9302)	44532 (13843)	43410 (11575)
YLC	42583 (8291)	47045 (9742)	46129 (17050)	42987 (9139)
All Groups	42057 (9583)	44843 (9176)	44554 (13611)	42188 (9047)

Firstly, we plotted the results graphically by group in order to determine any emerging response patterns (Figure 7-7).

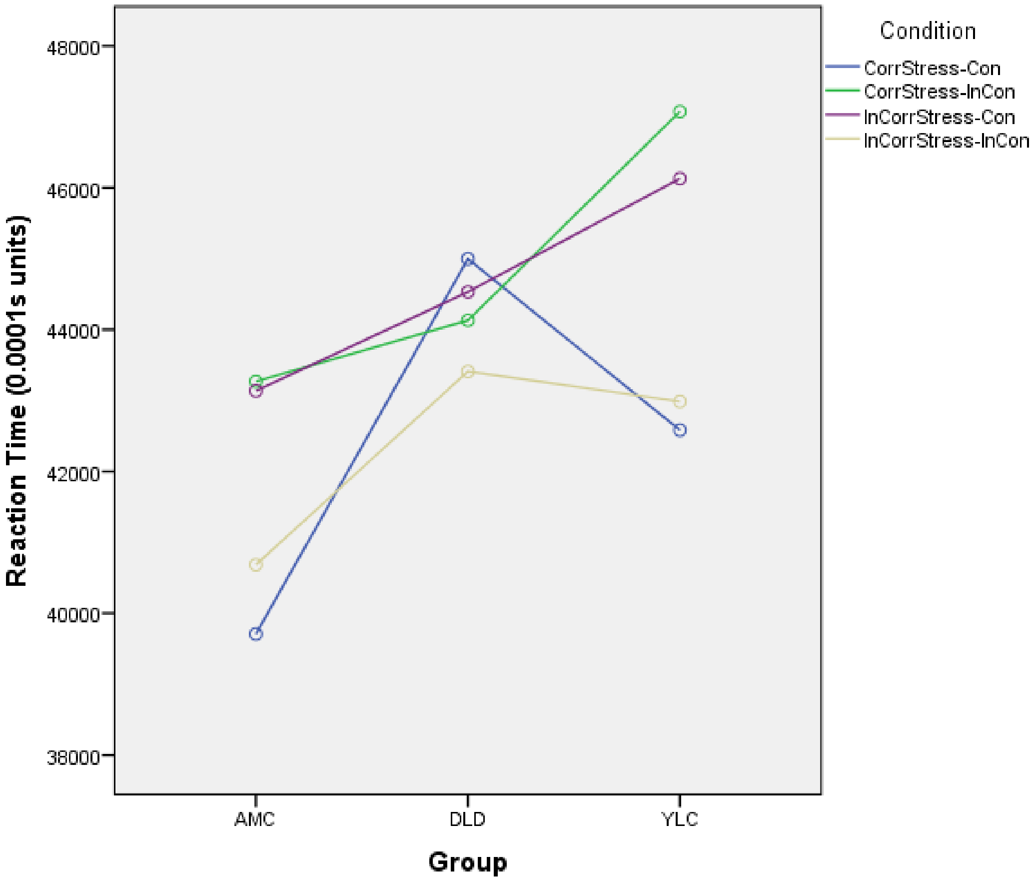


Figure 7-7 Graph showing the mean Reaction Times for each group by Condition

The resulting graph (Figure 7-7) shows an interesting between-group difference in the spread of reaction times. For the AMC and YLC groups, the CorrStress-Con and InCorrStress-InCon stimuli (blue and beige lines) appear to result in faster responses than the CorrStress-InCon and InCorrStress-Con stimuli (green and purple lines). This pairing is initially puzzling as it does not correspond to our previous categories of interest: Congruence of rhythm or correctness of the target.

To explore the data further, we performed a 3 x 2 x 2 repeated-measures ANOVA (Group – [AMC, DLD, YLC] x Rhythmic Congruence – [Congruent, Incongruent] x TargetStress – [Correct, Incorrect]. We had expected that Congruence of Rhythm would result in faster RTs, however there was no significant effect of Congruence (1, 53) = .009, *p* = .927. We had also expected that targets with Correct Stress would elicit faster responses than those with Incorrect Stress, however again there was no significant effect of TargetStress *F*(1, 53) = .037, *p* = .849. There was, however, a significant interaction between Congruence and TargetStress *F*(1, 53) = 11.925, *p* = .001. This interaction supports the visual pairings that can be seen for the AMC and YLC groups between CorrStress-Con and InCorrStress-InCon (blue and beige lines) and CorrStress-InCon and InCorrStress-Con (green and purple lines).

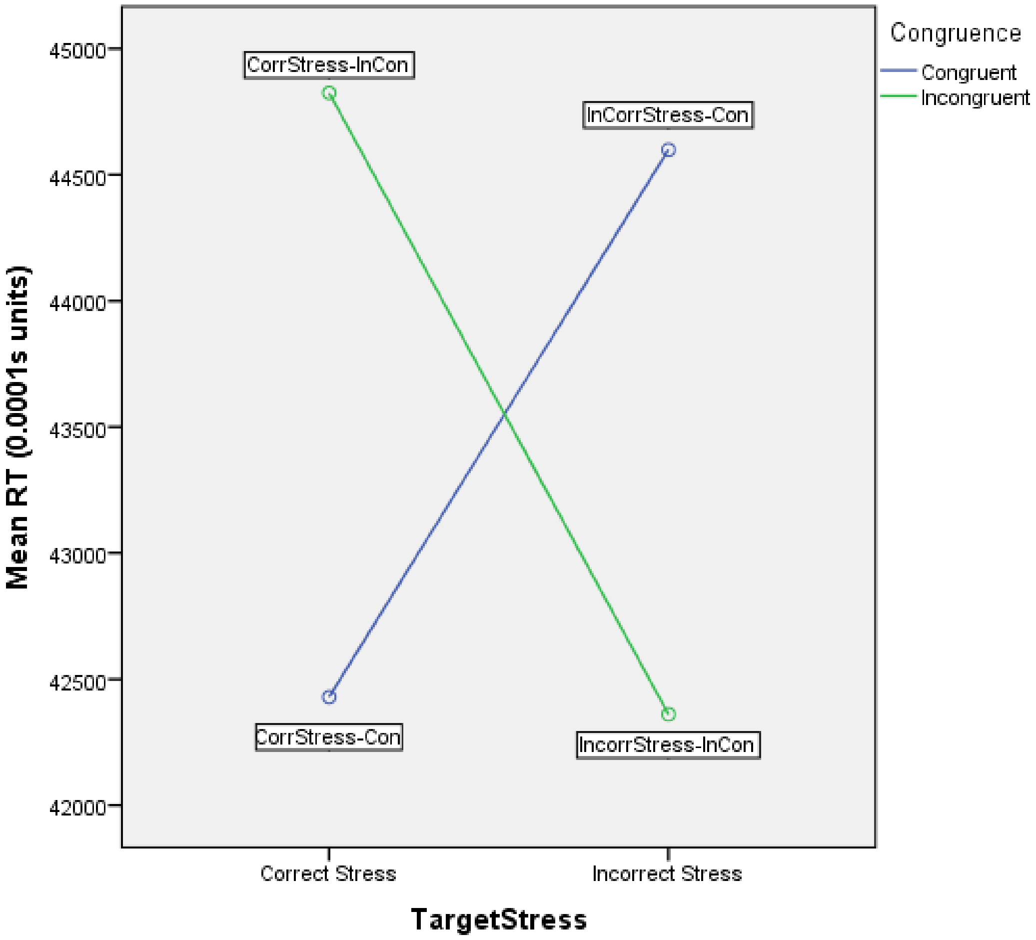


*Figure 7-8 Graph showing the Congruence*TargetStress Interaction*

In order to account for this unexpected result, we returned to a consideration of the stimuli types to see if there was a variable that would explain this. Remember that there were two contrasting rhythmic patterns in the carrier phrases. The CorrStress-Con and InCorrStress-InCon phrases consisted of a Sw trochaic rhythmic pattern (‘Jack’ sentences), whilst the InCorrStress-Con and CorrStress-InCon phrases consisted of a wS iambic rhythmic pattern (‘the boy’ sentences) – see Table 7-8.

Table 7-8 Example sentences for the four Conditions illustrating the Sw and wS carrier phrases

		Abbreviation	Stress pattern (sentence)										
Stress	**Rhythm**		**w**	**S**	**w**	**S**	**w**	**S**	**w**	**S**			
Sw	Trochaic	**CorrStress-Con**		JACK	is	REA	ding	BOOKS	a	BOUT	a	**CHI**	cken
Sw	Trochaic	**InCorrStress-InCon**		JACK	is	REA	ding	BOOKS	a	BOUT	a	chi	**CKEN**
wS	Iambic	**CorrStress-InCon**	The	BOY	is	REA	ding	BOOKS	a	BOUT	**CHI**	ckens	
wS	Iambic	**InCorrStress-Con**	The	BOY	is	REA	ding	BOOKS	a	BOUT	chi	**CKENS**	

Note: Sw sentences highlighted in pink

This prosodic element that the pairings have in common suggests that the interaction effect between Congruence and TargetStress may be caused by this third factor of the specific rhythmic alternation used in the carrier phrase, with Sw sequences resulting in faster reaction times than wS sentences, regardless of the Congruence of the concluding rhythm or the correct stressing of the target.

When the repeated-measures ANOVA was conducted for each group separately (2 x 2 - Congruence x TargetStress – DV - RT), there remained a significant Congruence * TargetStress interaction for the AMC group *F*(1,21) = 25.497, *p* = .000, and the YLC group *F*(1, 19) = 5.693, *p* = .028, but no significant interaction for the DLD group *F*(1, 13) = .020, *p* = .891. This suggests that the facilitating effect of Sw sentences was not a significant factor in the Reaction Times of the DLD group.

In order to explore this prosodic hypothesis further, two variables were created by taking the mean RT of CorrStress-Con and InCorrStress-InCon (SwCarrier) and of InCorrStress-Con and CorrStress-InCon (wSCarrier). A repeated-measures ANOVA (3 x 2 – Group [AMC, DLD, YLC] x Rhythm [SwCarrier, wSCarrier]; DV – RT) was conducted, then also for each group individually (1 x 2 (Rhythm)).

Table 7-9 Results of repeated-measures ANOVA for Carrier Phrase Rhythm by group

Reaction Times	**SwCarrier Mean (SD)**	**wSCarrier Mean (SD)**	**df**	***F***	***p***
AMC[a]	40196 (6619)	43202 (8730)	1, 21	25.497	.000
DLD	44204 (12561)	44331 (11383)	1, 13	.020	.891
YLC[a]	42785 (8034)	46602 (11567)	1, 19	5.693	.028
All Groups[a]	42123 (8892)	44699 (10401)	1, 53	11.925	.001

Note: a) SwCarrier < wSCarrier

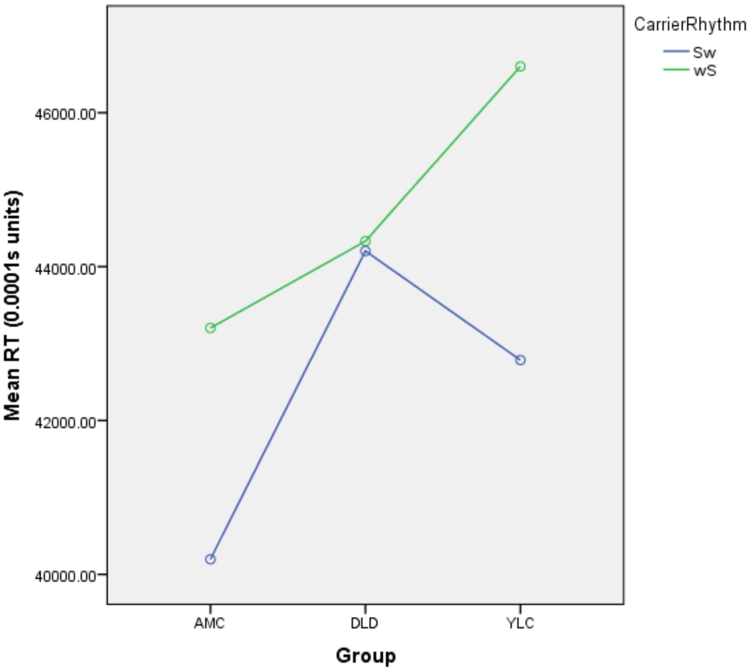


Figure 7-9 Graph of the RTs for CarrierRhythms (Sw and wS) by Group

There was a significant effect of Carrier Phrase rhythm on the RTs of the AMC and YLC group with Sw Carrier phrases responded to more quickly than wS Carrier phrases. There was no effect of Carrier Phrase Rhythm on the RTs of the DLD group.

We had expected that Congruent Rhythms and Correct Stress in the target word would result in faster RTs, however there was no significant effect of either Congruence or TargetStress on RTs for any of the groups. There was, however, a significant effect of CarrierPhrase Rhythm for the AMC and

YLC groups, with Sw CarrierPhrases resulting in significantly faster RTs than wS CarrierPhrases. There was no effect on RT for the DLD group for any of the rhythmic manipulations.

7.4 Relationship with Acoustic Thresholds

We predicted that higher rates of accuracy in judging lexical stress would be associated with greater sensitivity to the acoustic dimensions of RiseTime, Frequency and Duration.

In order to investigate this, we conducted a series of correlations between the four Acoustic Threshold tasks (RiseTime, Frequency, Duration and Intensity) and Lexical Stress Score (Table 7-10)

Table 7-10 Correlations (Pearson one-tailed) for AMC and DLD groups between Acoustic Threshold and Lexical Stress Score

Task	Duration	Frequency	Intensity	Score
Rise Time	.322*	.793***	-.216	-.438**
Duration	-	.385*	.064	-.349*
Frequency	-	-	-.196	-.473**
Intensity	-	-	-	.046

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

Score correlated significantly with RiseTime (*p* = .004) and Frequency (*p* = .002) as well as with Duration (*p* = .018). The negative correlation indicates that the greater the acoustic sensitivity, the more accurate were judgments of lexical stress. There was no significant correlation with Intensity.

There was therefore a significant correlation between acoustic sensitivity to Rise Time, Frequency and Duration and successful judgements of lexical stress.

A set of regression equations was conducted to explore the unique variance in Lexical Stress performance accounted for by the predictors of Age, NVIQ and each AT measure in turn. Step 1 was always entered as Age (months), Step 2 as NVIQ and Step 3 as the AT measure. The dependent variable was the overall accuracy score.

Table 7-11 Results of Regressions exploring the unique variance in Lexical Stress score accounted for by Age (months), NVIQ and AT measures (AMC and DLD groups)

	b	**SEb**	**β**	***t***	***p***	**ΔR^2**	***p***
Model 1							
Age	.088	.027	.464	3.214	.003	.336	.000
NVIQ	.199	.213	.134	.934	.358	.040	.164
Rise Time	-.010	.006	-.265	-1.778	.085	.058	.085
Model 2							
Age	.096	.027	.505	3.481	.002	.336	.000
NVIQ	.242	.217	.163	1.115	.273	.040	.164
Duration	-.011	.010	-.169	-1.145	.261	.025	.261
Model 3							
Age	.084	.029	.444	2.931	.007	.347	.000
NVIQ	.031	.238	.021	.130	.897	.024	.294
Frequency	-1.769	.841	-.347	-2.104	.044	.083	.044
Model 4							
Age	.102	.029	.537	3.560	.001	.333	.000
NVIQ	.300	.220	.202	1.366	.182	.039	.173
Intensity	-.022	.315	-.010	-.069	.945	.000	.945

Age was a significant predictor of Lexical Stress Score in the models, explaining between 33.3% and 34.7% of unique variance, whereas NVIQ did not contribute significantly (*p* range .164 - .294) when entered as Step 2. Frequency contributed further significant unique variance (8.3%, *p* = .044) whilst the contribution of Rise Time (5.8%) did not reach significance (*p* = .085). Neither Duration nor Intensity contributed significantly when entered as Step 3 (*p* = .261, .945 respectively).

7.5 Effect of Entrainment

7.5.1 Effect of Entrainment on Scores

A 2 x 2 repeated-measures ANOVA (Group [AMC, DLD] x Entrainment [Entrained, Unentrained] – DV Score) was conducted. There was no significant effect of Entrainment *F*(1, 34) = .796, *p* = .379, and Entrainment did not interact with group *F*(1, 34) = 1.128, *p* = .296.

To see whether any effect of Entrainment differed according to the preceding rhythm, a further 2 x 2 x 2 repeated-measures ANOVA was conducted (2 x 2 x 2 – Group [AMC, DLD] x Entrainment

[Unentrained, Entrained] x Congruence [Congruent, Incongruent] – DV[Score]). There was no significant main effect of Entrainment *F*(1, 34) = .796, *p* = .379, and Entrainment did not significantly interact with Congruence *F*(1, 34) = .017, *p* = .896.

We further investigated whether any effect of Entrainment differed according to whether the target word was correctly or incorrectly stressed (2 x 2 x 2 – Group [AMC, DLD] x Entrainment [Unentrained, Entrained] x TargetStress [CorrectlyStressed, IncorrectlyStressed] – DV[Score]). Entrainment did not significantly interact with TargetStress *F*(1, 34) = .905, *p* =.348.

Finally, we investigated whether any effect of Entrainment differed between Trochaic and Iambic CarrierPhrases (2 x 2 x 2 – Group [AMC, DLD] x Entrainment [Unentrained, Entrained] x CarrierPhrase [SwCarrier, wSCarrier] – DV [Score]). There was no significant interaction of Entrainment with CarrierPhrase *F*(1, 34) = .908, *p* = .347.

Listening to an entraining beat therefore did not have a significant effect on the accuracy of children's judgements of lexical stress, regardless of the preceding rhythmic context. Nor did entrainment make a difference according to whether the target word was correctly or incorrectly stressed or according to the rhythmic pattern of the carrier phrase.

7.5.2 Effect of Entrainment on Reaction Times

We also investigated whether an entraining beat affected Reaction Times. To do this we initially conducted a repeated-measures ANOVA (2 x 2 – Group [AMC, DLD] x Entrainment [Unentrained, Entrained] – DV[Reaction Time]. There was no significant effect of Entrainment *F*(1, 34) = 2.942, *p* = .095, and nor did Entrainment interact with Group *F*(1, 34) = .246, *p* = .623.

We further investigated any differential effect for Congruence of the final word with the preceding rhythm with a repeated-measures ANOVA (2 x 2 x 2 - Group[AMC,DLD] x Entrainment (Unentrained, Entrained) x Congruence [Congruent, Incongruent] – DV[RT]. There was no significant interaction of Entrainment with Congruence *F*(1, 34) = 3.161, *p* = .084.

We also investigated whether Entrainment differentially affected RT when words were correctly or incorrectly stressed using a repeated-measures ANOVA (2 x 2 x 2 – Group [AMC,DLD] x Entrainment (Unentrained, Entrained) x TargetStress [CorrectlyStressed, IncorrectlyStressed] – DV[RT]. There was no significant interaction of Entrainment with TargetStress *F*(1, 34) = .855, *p* = .362.

Finally, we investigated whether any Entrainment effects on RT differed according to CarrierPhrase rhythm (trochaic or iambic). We used a repeated-measures ANOVA (2 x 2 x 2 – Group [AMC,DLD] x

Entrainment (Unentrained, Entrained) x CarrierPhrase [SwCarrier, wSCarrier] – DV[RT]. There was no significant interaction of Entrainment with CarrierPhrase *F*(1, 34) = 1.956, *p* = .171.

Listening to an entraining beat therefore did not have an effect on the time children took to make their judgements of lexical stress, with no differential effect found for either the rhythmic context of the preceding sentence, whether the target word was correctly or incorrectly stressed or for the rhythmic pattern of the carrier phrase.

7.6 Summary of Results

Results were analysed according to two outcome measures – Accuracy and Reaction Time.

Taking the accuracy data first, there was a significant effect of Congruence for AMC and DLD groups, with words whose rhythmic pattern was congruent with the preceding carrier phrase being judged more accurately than those whose rhythmic pattern disrupted the carrier phrase rhythm. This effect was not observed for the YLC group. There was also a significant effect of TargetStress on accuracy for all groups, with target words which were correctly stressed being judged more accurately than those which were incorrectly stressed. This meant that all children were more likely to accept an incorrectly stressed word (e.g. chiCKEN) as correct, than to reject correctly stressed words (e.g. CHIcken). The effect of this was numerically greatest for the YLC group.

For the Reaction Time data, there was no significant effect of Congruence or of TargetStress for any of the participant groups. Conversely, we found a significant interaction between RTs for Congruence and TargetStress. This seems to have been driven by the specific rhythm used in the CarrierPhrase, with Sw CarrierPhrases resulting in faster reactions than wS CarrierPhrases, regardless of whether the final word was congruent or not, or correctly stressed or not. The DLD group did not differ in RT between Sw and wS CarrierPhrases.

Lexical stress score correlated significantly with acoustic sensitivity to Rise Time, Frequency and Duration, whilst there was no significant association with Intensity. Regressions indicated that Age was a significant contributor to score, with sensitivity to Frequency adding significant additional variance.

There was no significant effect of entrainment on either accuracy or RT for any of the groups or variables investigated.

7.7 Discussion

This experiment was designed to investigate whether the rhythmic context provided by the preceding sentence exerted an influence on children's abilities to judge whether a target word was correctly or incorrectly stressed. Children listened to sentences in which the final word either continued the preceding rhythm or disrupted it. This final word was then either correctly stressed or incorrectly stressed, creating four different Conditions: CorrStress-Con, CorrStress-InCon, InCorrStress-Con and InCorrStress-InCon. The scores and reaction times were then analysed to examine the effect of rhythmic context on the results.

Based on our work investigating representations of lexical stress patterns of words presented in isolation (Richards & Goswami, 2015), we predicted that the children with DLD would be less able to judge whether the final word of the sequence was produced correctly. Our prediction of a DLD deficit in accuracy was only partially borne out by the data. Whilst the DLD group mean was consistently lower than that of the AMC group (see Table 7-4), this difference proved not to be significant. However, the presence of a consistent trend suggests that this task was nevertheless more challenging for this group than for their age-matched peers. The reason for the lack of a more marked group difference in this sample compared to the previous experiment could be due to the different task demands. In our original task, children heard two words and had to decide which of the two was correct. This meant that they had to retain and compare the two stimuli in working memory whilst comparing them against their stored representation. In the present task, they had only one word to compare. This reduction in memory load may have led to a better performance from the DLD group thereby narrowing the gap between them and their age-matched peers. It could also be that the use of a linguistic context instead of an isolated presentation facilitated the lexical judgement aspect of the task.

The first variable of interest was that of Congruence of the final word with the rhythm of the preceding Carrier Phrase. We found that for AMC and DLD children, words whose rhythm was congruent with that of the preceding phrase were judged more accurately than those whose rhythm was incongruent. This result is in line with theories of temporal expectancies. If the preceding rhythm directs attentional resources so that they are maximised to the most salient point in time, then judgements of stimuli arriving at that timepoint should be more accurate than those which arrive unexpectedly. Of particular importance here is that the syllable rate was constant, whilst the timing of the final stressed syllable altered between congruent and incongruent variables. When the stressed syllable arrived at the expected timepoint, judgement was facilitated. The arrival of an unstressed syllable in the expected timeslot (incongruent condition) did not facilitate judgement. This suggests that it is the temporal location of stressed syllables in the previous rhythm which is the driving force for the allocation of attentional resources, and the efficient processing of stressed syllables which drives lexical stress judgements. The DLD children also showed a beneficial effect for Congruence, which suggests that when the preceding rhythm is strong and consistent (as it was in our stimuli) the temporal coincidence of the expected timepoint of the target stressed syllable and its actual arrival facilitates the decision-making process of children with DLD in judging whether presented stress patterns are accurate.

Based on Dilley & McAuley (2008)'s work on distal prosodic effects, we considered whether the Congruence effect for the CorrStress stimuli could result from internally generated expectations for the rhythmic pattern of the final word. They had found that a strong Sw pattern preceding neutrally stressed morphemes caused adults to infer a Sw pattern for the final word. If a similar effect obtained for our children, then it could have caused the CorrStress-InCon target (CHIcken) to be misheard as (chiCKEN) because this would have conformed to previous expectations. For InCorrStress targets, a distal prosodic effect would have caused InCorrStress-Con targets (chiCKEN) to be misheard as (CHIcken) and thus wrongly judged as 'correct', leading to lower accuracy for InCorrStress-Con than InCorrStress-InCon targets. This was not the result obtained as InCorrStress-Con words were judged more accurately than InCorrStress-InCon words. It seems that the effect of temporal expectancy is therefore the more likely explanation for the result pattern obtained.

The second variable of interest was that of the TargetStress of the final word. We found that this was a significant factor in judgement accuracy for all of our participant groups, with all groups having a tendency to accept incorrectly stressed items (e.g. chiCKEN) as correct. This finding was strongest for the YLC group and weakest for the AMC group, suggesting that although individual stress patterns begin to be established early in the word-learning process (Curtin et al. 2005), the robustness of representation for individual lexical items continues to be strengthened over the course of development. The results for the DLD group fell between these two extremes which suggests that these children are capable of developing accurate representations of lexical stress, but are doing so at a slower rate than their typically developing peers.

Consideration of Reaction Times revealed a further rhythmic factor. For RTs, neither Congruence not TargetStress had a significant impact, however the specific rhythm of the CarrierPhrase had a significant impact on how quickly children responded. When the CarrierPhrase was Sw ('Jack' sentences) children responded faster than when the CarrierPhrase was wS ('the boy' sentences). There is evidence from various sources that Sw structures are privileged in English – infants establish this pattern earlier than wS patterns (Jusczyk, Houston, & Newsome, 1999), Sw structures dominate spoken English (Cutler & Carter, 1987) and Gerken (1996) suggests that children's own productions are dominated by a Sw template. Here, the Sw rhythm resulted in children responding faster (but not more accurately) than in the wS conditions. This cannot be explained by general temporal expectancies, since these should operate across all stressed syllables regardless of the internal patterning. A possible explanation is that since Sw structures are so dominant in English, language processing happens more efficiently (i.e. quickly) for stimuli which conform to this predominant pattern. The the language system is primed, developmentally to expect Sw rhythms, while wS rhythms are more anomalous. If we propose that this effect is driven by a Sw preference bias, then the lack of an RT effect for the DLD children may imply that they have not yet established a Sw preference.

We had hypothesised that listening to an entraining beat would induce a stronger temporal expectancy for children, thereby enhancing any rhythmic effects found in the Unentrained condition and potentially leading to faster decision-making. However our results indicated that entrainment had no significant effects on either accuracy or reaction times for any of the groups. We saw a potential effect of temporal expectancy in the Unentrained condition, and so it could be that our carrier phrases were sufficiently rhythmic on their own to induce a strong temporal expectancy, rendering the additional entrainment phase redundant.

Increased sensitivity to the acoustic cues of Rise Time, Frequency and Duration was associated with more accurate judgements of lexical stress, however the regression analysis indicates that some of this correlation was mediated by age. Frequency remained a significant predictor of Score, however, once age was accounted for, whilst Rise Time contributed the next greatest amount of variance. These results indicate that proficiency in lexical stress judgments improves with age, but also that acoustic sensitivity plays an additional contributory role.

The results of this experiment demonstrate that the rhythmic context preceding a target word has a significant impact on lexical stress judgements. Words whose rhythmic pattern was congruent with the preceding rhythm were judged more accurately by AMC and DLD children than those whose pattern was incongruent, suggesting a potential role for temporal expectancies at the stressed syllable rate in facilitating lexical stress judgements. The target stress pattern also had a significant impact on accuracy, with all children more likely to accept incorrectly stressed words than reject correctly stressed ones. This effect was greatest for the YLC children, suggesting that the consolidation of stress patterns for individual lexical items is an ongoing process, with representations being gradually refined and strengthened over the course of development. The rhythm of the preceding CarrierPhrase also impacted on reaction times for AMC and YLC children, with both groups responding faster (but not more accurately) to CarrierPhrases in which the rhythm was Sw than they did to phrases which were wS. The DLD group did not respond differently to the different CarrierPhrases. This suggests that typically-developing children have more efficient processing of Sw stimuli, but that the DLD group have not developed the same Sw processing bias.

8 Experimental Task 4 - Sentence Repetition

Moving on from rhythm operating at the word level, we were interested to investigate how rhythmic factors may influence representations the phrase and sentence levels. We saw in Section 1.4.4 that prosodic cues might serve as a preliminary form of chunking of larger syntactic units, enabling more efficient processing to take place. A task frequently used to assess sentence-level aspects of children's language is Sentence Repetition. We therefore decided to explore the influence of rhythm in the performance of sentence repetition.

The following section will firstly explore what is known about sentence repetition in general, before considering the specific role that rhythm might play in aspects of the sentence repetition process.

8.1 Theoretical Basis of the Task

8.1.1 Sentence Repetition, Language and Language Disorder

Sentence Repetition is a task frequently included in language assessments (for example, it is a sub-test in each incarnation of the CELF). It typically involves the assessor reading out a series of increasingly complex sentences which the child has to repeat verbatim. Children with DLD often perform particularly poorly in the repetition task, so much so that Conti-Ramsden, Botting, & Faragher (2001) posited poor performance in sentence repetition as a potential psycholinguistic marker for DLD. In their study of 155 11-year-olds with DLD, they found that performance in the CELF-R Recalling Sentences task was the marker task with the greatest accuracy in identifying both children with current and those with resolved language difficulties with a specificity of 90% and a sensitivity of 85%. In a further study with 3 and 4 year-olds, performance in sentence repetition was again found to be the best marker of expressive language delay (Everitt, Hannaford & Conti-Ramsden, 2013), whilst Archibald & Joanisse (2009) found that sentence repetition was useful in identifying children with language and combined language and working memory impairments. Alloway & Gathercole (2005) found that sentence repetition accounted for a small, but significant, percentage of unique variance in language skills in their group of participants with special educational needs as well as contributing significant variance to measures of reading. Poor sentence repetition skills are also found in children who struggle to comprehend written text (Marshall & Nation, 2003). These children also had weak oral language skills and so may have some overlap with children typically classified as having DLD. Sentence Repetition has also been investigated in a variety of other languages such as Cantonese (Stokes, Wong, Fletcher, & Leonard, 2006), Czech (Smólik & Vávrů, 2013) and Italian (Devescovi & Caselli, 2007) and has been found to be a useful clinical marker of language difficulty as well as a measure of language development.

Difficulty in sentence repetition is therefore a hallmark of DLD, however, despite its widespread use as a clinical tool, the mechanisms underlying task difficulty and success remain the subject of debate.

8.1.2 Sentence Repetition and Memory systems

As with NWR, a frequently adopted means of understanding the task demands of sentence repetition is to focus on memory systems. This is where Conti-Ramsden et al. (2001) locate the core deficit resulting in difficulties in sentence repetition, suggesting either a limited short-term memory capacity or a rapid decay of the memory trace as being possible sources of repetition errors. They posit that these memory factors could explain the difficulties typically seen in repetition tasks at both the nonword and sentence level.

Phonological short-term memory has been implicated in the repetition of individual (non)words (Gathercole & Baddeley, 1990) and researchers investigating sentence repetition have also found that the task has a phonological capacity component. As the length and number of words in a sentence increase, there is a significant negative impact on accuracy of recall (Willis & Gathercole, 2001). This implies that children with shorter spans will recall less material accurately.

Sentence Repetition does not solely rely on phonological short-term memory, however. Rohl & Pratt (1995) found that sentence repetition seemed to be tapping a separate factor from their tests of phonological working memory. Marshall & Nation (2003) found that their group of ‘poor comprehenders’ (children who struggle to follow written text) had poor sentence repetition scores but performed age-appropriately on tests of short-term memory. Similarly, Alloway, Gathercole, Willis & Adams (2004) found that their phonological loop measures could not account for the sentence repetition scores they obtained.

In Chapter 6 (Nonword Repetition), we saw that on-line and long-term memory systems combine in a process of redintegration to enable successful repetition. A similar process is thought to occur in Sentence Repetition, with prior phonological, semantic and syntactic knowledge all contributing to task success. This possibility was briefly discussed by Conti-Ramsden et al. (2001). They found that performance in Sentence Repetition tasks correlated significantly with other linguistic tasks presented (such as use of tense morphology) to a greater degree than the nonword repetition task, and interpreted this as indicating a greater linguistic processing load for the sentence-based tasks. The notion that sentence repetition involves systems other than short-term memory was also investigated by Potter & Lombardi (1990) and Lombardi & Potter (1992). In their experiments, they demonstrated that sentence recall was influenced by intrusion of priorly activated lexical items consistent with the overall meaning of the sentence. They used this evidence to conclude that sentences were regenerated in immediate recall from a representation of meaning, rather than purely from a short-term memory verbatim trace. They proposed that the recall of sentences was based on a propositional representation of meaning for which recently activated lexical items were chosen to fill an appropriate syntactic frame.

A further implication of the redintegration theory is that sentence repetition will also reflect syntactic competence. In fact, successful repetition of target syntactic structures has frequently been regarded as a measure of expressive morphosyntax (e.g, Devescovi & Caselli, 2007; Riches, 2012; Seeff-Gabriel, Chiat, & Dodd, 2010).

Success in Sentence Repetition can therefore be regarded as the product not of a single aspect of language processing, be that phonological short-term memory, morphology or syntax, but as having the potential to reflect impairment across a range of linguistic processes (Klem et al., 2015; Riches, 2012). If so, then one such previously unexplored linguistic influence could be in the area of rhythm.

8.1.3 Sentence Repetition and Rhythm

Rhythm is rarely discussed in the context of Sentence Repetition tasks, however there is reason to propose a role for rhythmic processing in each of the skill areas outlined above. We proposed in the discussion of nonword repetition that the prosodic structure of a target word could provide a scaffolding framework in phonological short-term memory, facilitating subsequent repetition of that target. A similar effect could obtain for the repetition of novel sentences. If we view prosodic or rhythmic representation as “the essential skeleton that holds different syllables together” (Frazier, Carlson, & Clifton, 2006, p244), then rhythm will be operating to create a scaffolding framework at the sentence, as well as lexical, level. In this view, children who are able to form a representation of the incoming rhythmic template will be at an advantage in retaining the more detailed phonological content of the sentence. Furthermore, if rhythm enables chunking of sentences into more manageable units, then this may result in more efficient processing of the component semantic and syntactic features of each unit. Conversely, absent or imprecise rhythmic chunking may lead to attempts to represent longer units of speech, creating more unwieldy amounts of material for the language system to process.

If Sentence Repetition also calls upon long-term knowledge of linguistic structures, then rhythmic sensitivity may also play a role here. Learning of individual lexical items is thought to be dependent on phonological short-term memory (e.g. Gathercole, 2006; Archibald & Gathercole, 2007), of which the prosodic structure is part. As discussed in Nonword Repetition (Chapter 6) poor sensitivity to prosodic structure may lead to erroneous or underspecified representations of individual words, thus diminishing the ability to call on these words from long-term storage in a Sentence Repetition task. We also saw in Section 1.4.4 that prosody has a role to play in the development of grammatical structures, and in Section 1.3.2 that poor sensitivity to rhythmic cues could potentially impact on the ability to extract grammatical information from speech. If a child has impoverished knowledge of the underlying syntactic and grammatical structures, then they will be less able to draw on these resources effectively than a child for whom such knowledge is secure. Furthermore, rhythm's role in long-term knowledge may not be just a developmental one (enabling the original creation of robust syntactic structures), but it may also be that rhythmic templates at the word, phrase and clausal level are part of the long-term knowledge that facilitates redintegration. Speakers exploit the typical rhythmic patterns of their language in order to segment new words and phrases (Cutler, Mehler, Norris, & Segui, 1992) and so it seems likely that expectations of typical rhythmic patterns also form part of the long-term linguistic knowledge called upon in the redintegration process. If a child has good sensitivity to rhythm, any rhythmic expectancies created will be both varied (in terms of the child having developed knowledge of a wide range of potential structures) and robust, providing a strong potential source of bootstrapping for task completion. Conversely, children with poorer rhythmic sensitivity are more likely to have weak or ill-defined rhythmic templates, more limited in scope, which will provide less support in attempting to complete the task.

The presence of rhythm has been found to have a beneficial role in sentence repetition tasks. Polišenská, Chiat, & Roy (2015) contrasted repetition of sentences which had been produced with natural prosody with those produced as word strings in a study with typically developing 4-5 year olds in both Czech and English and found a significant beneficial effect of prosody. Consistent with developmental theories, they attributed this to the children benefiting from the prosodic structure to chunk the incoming sentence more effectively for subsequent recall.

Less is known about the role of specific rhythmic patterns in facilitating recall, however there have been some interesting studies focusing on the circumstances facilitating or hindering the production of content and function words in specific rhythmic contexts.

Function words are regarded as a 'closed class', and have a predominantly grammatical function, including prepositions, auxiliaries, copulas and conjunctions. Content words form an 'open class' of lexical items, with new terms readily integrated into the existing lexicon, and include nouns, verbs, adjectives and adverbs (Segalowitz & Lane, 2000). Assignment of category (content or function) is therefore based on the grammatical status of the target word. Many studies of repetition have used the division of target words into content and function (e.g. Polišenská et al., 2015; Seeff-Gabriel et al., 2010) and a general observation has been that young children tend to preserve content words, but are more likely to omit function words (e.g. Chiat & Roy, 2007; Gerken, Landau, & Remez, 1990; Scholes, 1970). A similar phenomenon has been noted with regard to function and content words in children's spontaneous speech, together with the observation that children with DLD have particular difficulty with function words such as articles and pronouns (McGregor & Leonard, 1994; Chiat & Hirson, 1987). Furthermore, children with weaker language skills are more likely to omit these words in repetition tasks (Gerken et al., 1990).

The division into content and function words also has implications for our understanding of the role of rhythm, since content words typically contain a stressed syllable, whereas function words and morphemes in English are typically unstressed. Stress is a predictor of repetition accuracy with children preserving more strong syllables than weak syllables (Gerken, 1994b), such that Gerken et al. (1990) concluded that 'the degree of stress a syllable receives affects the likelihood it will be preserved in imitative speech' (p207). Strongly stressed content syllables are therefore better repeated than unstressed function syllables. However, from our theoretical perspective, there is a potential confound in the data so far. In previous studies, content syllables (and only content syllables) were stressed whilst function syllables (and only function syllables) were unstressed. This allows for alternative interpretations of the data, particularly with regard to function syllables. Firstly, the explanation may be a grammatical one in which inaccurate repetition of a function syllable reflects poor morphosyntactic competence (which poor sentence repetition performance is thought to reflect). Such an account would sit well with domain-specific accounts of DLD which argue for specific deficits in grammatical processing (e.g. van der Lely & Christian, 2000; van der Lely, Rosen, & Adlard, 2004). Alternatively, the explanation may be a rhythmic one, in which the omission of a function syllable reflects its typical rhythmic status as a weak, unstressed syllable. This account would point more towards a role for rhythmic properties of language in determining some aspects of language output. Because grammatical and rhythmic properties were conflated in the target stimuli in prior work, we cannot reliably conclude which explanation is the most appropriate.

The influence of rhythmic structure may further extend beyond individual words and syllables to operate at larger levels of linguistic organisation. In a series of experiments, (Gerken, 1991, 1994a, 1994b, 1996) showed that metrical context influenced the production of object articles in 2-year-old children in a sentence repetition task. She used a prosodic phonology account to propose that children have a rhythmic (or metrical) template for producing language structures and that (at least in English) this follows a Sw pattern. If target syllables do not fit into this template then they are more likely to be omitted, explaining the observation that children were more likely to omit an article when it was not part of a Sw metrical foot. For example, take the two sentences 'he kicks *the* pig' (wS*w*S) and 'he catches *the* pig' (wSw*w*S). In the former sentence, 'the' occurs after the strong syllable 'kicks', thus forming a Sw unit and is preferentially retained, whereas in the latter sentence 'the' occurs after the weak syllable '-es', does not form part of a Sw foot (is 'unfooted') and so is more likely to be omitted.

McGregor & Leonard (1994) also studied repetition of phrases in slightly older (3-5 year old) children with DLD. They found that weakly stressed function words (e.g. he, you, the) occurring at the beginning of a phonological phrase were more frequently omitted than were strongly stressed proper nouns. The phrase-initial function words were also more poorly repeated than those occurring later in the sentence in a Sw context. They interpreted their findings within a framework which supported Gerken's 1994 work, suggesting the primacy of a strong-weak production template in young children's speech. Both of these studies therefore suggest a role for rhythmic context in determining accuracy of sentence repetition, with function words outside Sw contexts particularly vulnerable.

There are therefore several questions arising from an examination of the potential role of rhythm in sentence repetition tasks which we wanted to explore in designing our experiment.

Firstly, there is the potential role of the overall rhythmic structure of the target sentence. Based on the predictions of temporal expectancies and neural oscillations, we expected that sentences in which there was a regularly recurring rhythmic structure would be more accurately repeated than those in which the rhythm was more irregular. The impact of overall target rhythms in sentences has not previously been examined. Secondly, we expected that sensitivity to the acoustic cues to rhythm (i.e amplitude rise time) would relate to successful task completion. If so, this would indicate a role for rhythmic sensitivity in repetition tasks at the level of the sentence, a relationship which has not previously been investigated. Thirdly, we were interested to investigate directly the overlap between the grammatical status (content or function) and the stress status (stressed or unstressed) of target words. This confounding factor has not previously been systematically examined in repetition tasks. In addition, we wanted to further investigate the relationship proposed by Gerken (1996) and McGregor & Leonard (1994) between rhythmic context and accurate repetition of unstressed function words, by systematically manipulating the stress status of the syllable preceding a given function word. Examination of these various factors could provide valuable insights into whether rhythm forms part of the task parameters relevant to successful sentence repetition.

8.2 Designing the Task

This experiment was designed in order to allow the investigation of several hypotheses:

1) That sentences with a regular rhythm would be more accurately repeated than those with a variable rhythm
2) That sensitivity to acoustic cues (rise time, duration, frequency, intensity) would relate to accuracy of children's sentence repetition.
3) That function words carrying stress would be more accurately repeated than function words which were unstressed
4) That the immediate prosodic context would affect repetition of unstressed function words, with unstressed function words repeated more accurately following a strong syllable than when following a weak syllable

8.3 Constructing the Stimuli

In order to create sentences that could provide relevant data for these different hypotheses, there were several considerations that had to be taken into account:

1) Syntactic structure
2) Sentence Rhythm
3) Prosodic context of unstressed function words
4) Stress patterns and Grammatical status of words

All of these were systematically considered and manipulated in the creation of the sentence stimuli.

8.3.1 Syntactic Structure

In choosing the target syntactic structures to be repeated, there was a trade-off between considerations of length and of complexity. The sentences needed to be short enough to be accessible and yet sufficiently complex to cause some difficulty to all participants in order to enable error analysis to take place.

Two target structures were chosen:

Structure 1 (Embedded Sentences) consisted of a 12-syllable sentence containing an embedded sentence. This structure was chosen because it is the final item of the TROG (Bishop, 1983) and therefore likely to cause difficulties for AMC children as well as the YLC and DLD groups. Each sentence followed the same basic structure (see Figure 8-1) with the lexical items varying in their target rhythm (see Section 8.3.4).

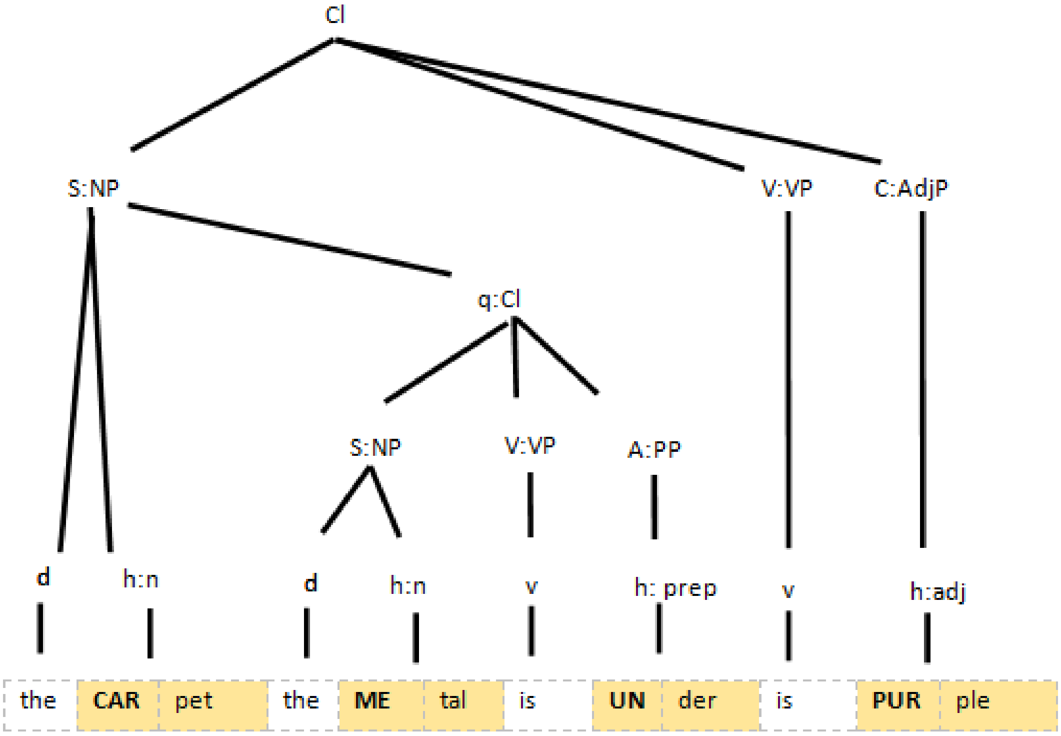


Figure 8-1 Figure illustrating the syntax of the Embedded Sentence structure. Yellow shading indicates the lexical items which varied between exemplars

Structure 2 (Complex Sentences) consisted of a 15-syllable sentence composed of a main clause and a subordinate clause introduced by a subordinating conjunction. This structure was chosen as clinical observations of children with DLD completing the Recalling Sentences task of the CELF suggested that this was the stage of complexity at which children tended to begin to struggle to complete the task accurately. As with the Embedded Sentences, the structure of the complex sentences remained the same whilst lexical items (which varied in stress patterns) varied between exemplars.

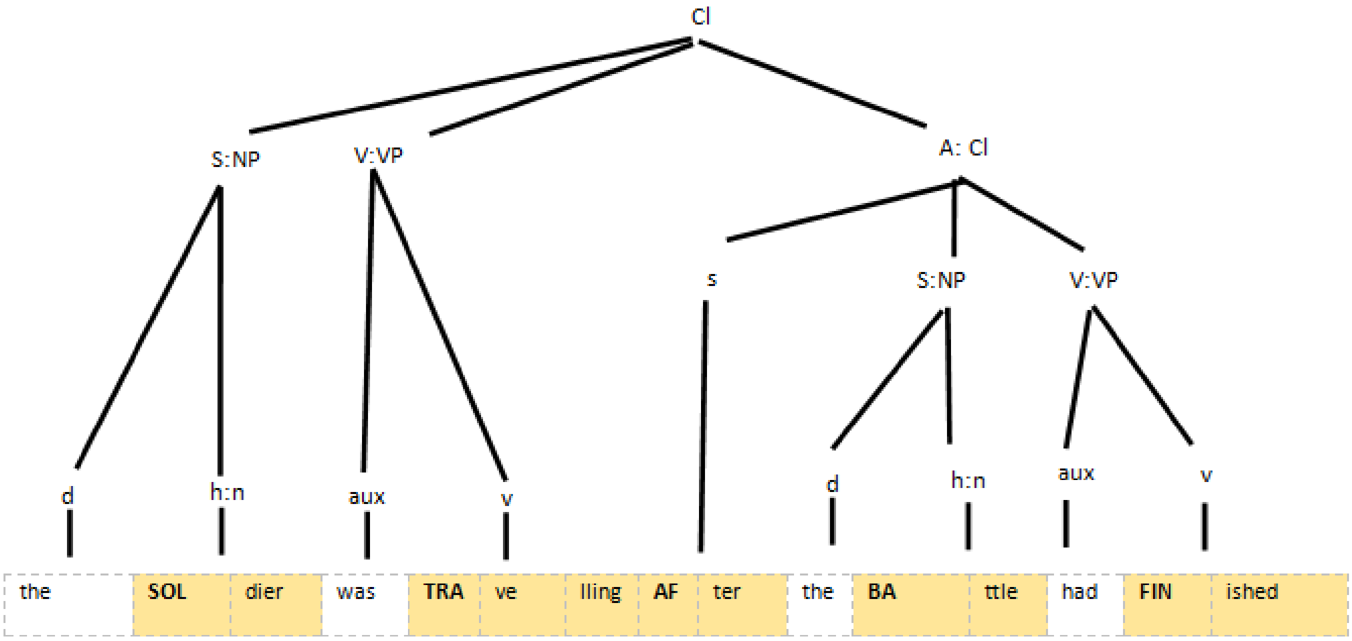


Figure 8-2 Figure illustrating the syntax of the Complex Sentence structure. Yellow shading indicates the lexical items which varied between exemplars

All lexical items within sentences (e.g. carpet, hotel) were matched for frequency across sentence types using figures from the British National Corpus (Leech, Rayson & Wilson, 2001).

8.3.2 Sentence Rhythm

Within each syntactic structure, two rhythmic patterns were created so that the rhythm of the resulting sentence was either regular and predictable (e.g. wwSwwSwwSwwS) (**Regular)** or variable and unpredictable (e.g. wSwwSwwwwSwwSwwS) **(Variable)**.

Table 8-1 Examples of Regular and Variable Rhythm Embedded Sentences

Type	**Example Sentence**											
Embedded Regular	the	**CAR**	pet	the	**ME**	tal	is	**UN**	der	is	**PUR**	ple
Embedded Variable	the	**TA**	bles	the	com	**PU**	ter	is	**UN**	der	are	**BIG**

Table 8-2 Examples of Regular and Variable Rhythm Complex Sentences

Complex Regular	the	**STU**	dents	were	**AR**	gu	ing	**AF**	ter	the	**MEE**	ting	had	**STAR**	ted
Complex Variable	the	**TEAM**	were	**CE**	le	bra	ting	be	**FORE**	the	**MATCH**	had	con	**CLU**	ded

The overall syntactic structure of each sentence-type remained the same.

8.3.3 Prosodic Context of unstressed function words

McGregor & Leonard (1994) hypothesised that function words were more likely to be retained if they occurred following strong syllables than following weak syllables. The sentences were therefore carefully constructed so that in half the sentences the function words (the, is, had, was) followed a strong syllable (Sw context) and in half they did not (ww context). Sw and ww context function words were equally distributed between regular and variable rhythms and embedded and complex sentence types.

Table 8-3 Example Sentences showing prosodic context of function words – Embedded Sentences

Type	**Example Sentence**											
Embedded Regular - ww	the	**CAR**	pet	the	**ME**	tal	is	**UN**	der	is	**PUR**	ple
Embedded Variable – ww	the	**TA**	bles	the	com	**PU**	ter	is	**UN**	der	are	**BIG**
Embedded Regular – Sw	the	dis	**PLAY**	the	gui	**TAR**	is	be	**HIND**	is	com	**PLETE**
Embedded Variable - Sw	the	ma	**CHINE**	the	ma	ga	**ZINE**	is	be	**HIND**	is	**SHORT**

Note: purple shading indicates function words in a ww context; yellow shading indicates function words in a Sw context

Table 8-4 Example Sentences showing prosodic context of function words – Complex Sentences

Type	**Example Sentences**														
Com-Reg –ww	the	**STU**	dents	were	**AR**	gu	ing	**AF**	ter	the	**MEE**	ting	had	**STAR**	ted
Com-Var – ww	the	ex	**PLO**	rer	was	**RES**	ting	**AF**	ter	the	ad	**VEN**	ture	had	**STOPPED**
Com-Reg – Sw	the	ho	**TEL**	was	co	**LLAP**	sing	be	**FORE**	the	po	**LICE**	had	pre	**PARED**
Com-Var - Sw	the	**TEAM**	were	**CE**	le	bra	ting	be	**FORE**	the	**MATCH**	had	con	**CLU**	ded

Note: purple shading indicates function words in a ww context; yellow shading indicates function words in a Sw context

8.3.4 Stress Patterns and Grammatical Status of words

Many of the content words also contained weak syllables in variety of prosodic patterns (e.g. TAble – Sw; hoTEL – wS; comPUter - wSw). These were chosen so as to comply with the overall rhythmic structure of the sentence as outlined in Section 8.3.2, and also so as to provide a range of different word level rhythms. This would enable the comparison of weak and strong syllables within content words.

As discussed in Section 8.1.3, we wanted to investigate whether stress status (stressed or unstressed) or grammatical status (function or content) was the major influence upon the result patterns found in previous studies.

In order to investigate this, each sentence contained a bisyllabic function word – either a preposition (e.g. UNder, BeHIND) or a conjunction (beFORE, AFter). As disyllables, these function words each contained an internal stress pattern (wS or Sw) and so enabled the examination of whether stress status (word with stressed syllable) interacted with grammatical status (function word).

There were therefore ultimately 8 different sentence types, systematically varying across three variables: Syntax (Embedded, Complex); Rhythm (Regular, Variable); Prosodic Context of function words (Sw, ww).

Table 8-5 Example Sentences of all four Embedded sentence-types

Type	**Example Sentence**											
Emb-Reg-ww	the	**CAR**	pet	the	**ME**	tal	is	**UN**	der	is	**PUR**	ple
Emb-Reg-Sw	the	dis	**PLAY**	the	gui	**TAR**	is	be	**HIND**	is	com	**PLETE**
Emb-Var-ww	the	**TA**	bles	the	com	**PU**	ter	is	**UN**	der	are	**BIG**
Emb-Var-Sw	the	ma	**CHINE**	the	ma	ga	**ZINE**	is	be	**HIND**	is	**SHORT**

Table 8-6 Example Sentences of all four Complex sentence-types

Type	**Example Sentences**														
Com - Reg- ww	the	**STU**	dents	were	**AR**	gu	ing	**AF**	ter	the	**MEE**	ting	had	**STAR**	ted
Com - Reg-- Sw	the	ho	**TEL**	was	co	**LLAP**	sing	be	**FORE**	the	po	**LICE**	had	pre	**PARED**
Com – Var- ww	the	ex	**PLO**	rer	was	**RES**	ting	**AF**	ter	the	ad	**VEN**	ture	had	**STOPPED**
Com – Var - Sw	the	**TEAM**	were	**CE**	le	bra	ting	be	**FORE**	the	**MATCH**	had	con	**CLU**	ded

8.4 Creating the stimuli

During recording, a beat of 750ms between stressed syllables was induced in the speaker via headphones using a metronome setting of 80bpm.

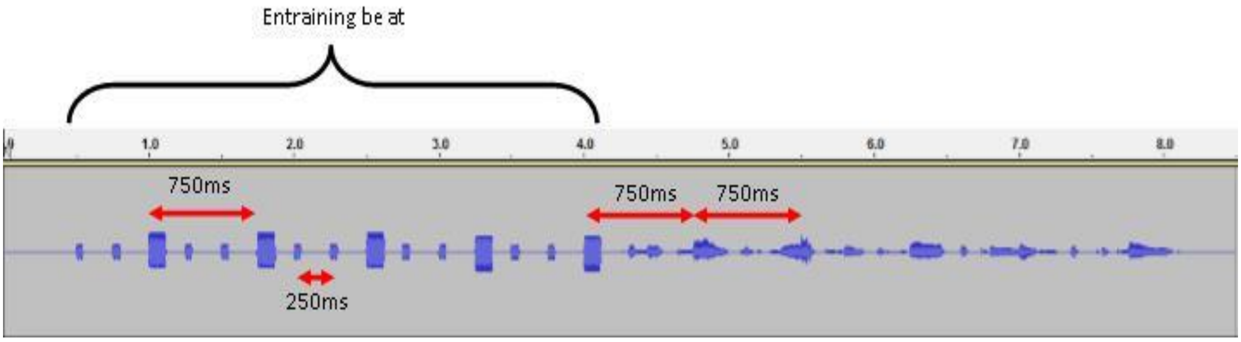


Figure 8-3 Soundwave of example entrained sentence 'the hotel was collapsing before the police had prepared'

Each sentence was spoken so that stressed syllables aligned with the pulse with a constant inter-stressed syllable interval of 750ms. The timing of this was verified and adjusted as necessary using Audacity software. The entraining beat also had an inter-stress interval of 750ms with 250ms intervals between the intervening weak beats (corresponding to unstressed syllables). The sentence followed the entraining beat such that the first stressed syllable of the stimulus was 750ms after the final strong entraining beat, thereby keeping the 750ms pulse constant. The number of inter-stress intervals in the entrainment section mirrored the number of inter-stress intervals in the stimulus.

8.5 Task Presentation

Two lists of 16 sentences were created with each list containing two each of the 8 sentence types above. Unentrained and Entrained versions of each list were created.

Sentences were presented via a laptop, using Presentation software with the children listening through headphones. The experimenter explained that they were going to hear some sentences and that their job was to repeat them as best they could, and to have a go even if they could not remember all of the words. For the entrained condition it was explained that they would hear some beeps first and that the beeps were there to help them hear the pattern of the sentence.

There were two practice sentences (one Embedded Sentence and one Complex Sentence) with the experimenter providing any additional instructions at this stage as necessary for understanding of the task. The children then proceeded with the experimental sentences with children’s responses recorded for later transcription.

A sample (*n* = 20) of recordings was independently transcribed by a qualified Speech and Language Therapist who was blind to the overall purpose of the task and to participant group. Inter-rater agreement was calculated for word score (96.99%) and syllable score (97.59%).

8.6 Preview of the Statistical Analysis

There were several questions which we wanted to explore through our analysis of this task. A summary of these is provided in the table below (Table 8-7).

Table 8-7 List of Questions addressed in this section together with our predictions

We thought Repetition may be affected by:	**We predicted that:**
The grammatical status of the target (Content or Function)	Content words would be more accurate than Function words
The stress status of the target (stressed or unstressed)	Stressed targets would be more accurate than Unstressed targets
The provision of an entraining rhythm	Entrained sentences would be more accurate than Unentrained sentences
The immediate prosodic context (Function word preceded by a stressed or unstressed syllable)	Function words in Sw contexts would be more accurate than in ww contexts
The regularity of sentence rhythm	Sentences with a regular rhythm would be more accurate than sentences with a variable rhythm
A relationship with acoustic sensitivity (rise time, duration, frequency, intensity)	Greater acoustic sensitivity would be related to better accuracy

Each of these questions will be addressed in turn in this statistical section. A further thematic discussion section will then attempt to draw these findings together into a holistic account of prosodic and grammatical influences on repetition accuracy.

8.7 Results: Global Accuracy

The children's responses were scored at several grains of analysis. At sentence level, responses were scored for the number of sentences repeated correctly in their entirety. At a smaller unit level, responses were scored for the number of words repeated correctly and within this, account was taken of whether the target word could be classed as a Function or a Content word. Reducing the grain size again, responses were scored for the number of syllables repeated correctly. The syllable responses were also coded as to whether the target syllable was stressed or unstressed and whether it occurred within a Function or Content word. This enabled a picture to emerge of the interplay of grammar and prosody across linguistic units.

Scores were calculated as the number of sentences, words or syllables repeated correctly in their entirety, expressed as a percentage of the total number of targets available. The maximum score at each level was therefore 100.

Table 8-8 Results of one-way ANOVAs comparing Group for overall accuracy scores

Score (% Targets Correct)	**AMC Mean (SD)**	**DLD Mean (SD)**	**YLC Mean (SD)**	***df***	***F***	***p***
Sentence Score^{a,b}	39.062 (22.364)	2.232 (5.262)	5.707 (8.821)	2, 37.764^{d}	29.216	.000
Word Score^{a,b}	82.108 (10.613)	46.796 (9.633)	54.188 (15.011)	2, 58	47.077	.000
Syllable Score^{a,b,c}	85.272 (9.556)	49.537 (9.914)	59.863 (15.755)	2,33.75^{d}	63.903	.000

Note: a)AMC > DLD b) AMC > YLC c) YLC > DLD d) Welch's *F* and *df* used due to significant Levene's test

A one-way ANOVA (DV- Score) was conducted for each grain-size. The AMC group scored significantly more highly than the DLD and YLC groups at each level of analysis (*p* = .000 throughout). The DLD and YLC groups did not differ from each other at either sentence or word level (*p* = .303 post-hoc Games-Howell, *p* = .243 respectively). At syllable level, however, the YLC group were marginally significantly more accurate than the DLD group (*p* = .05).

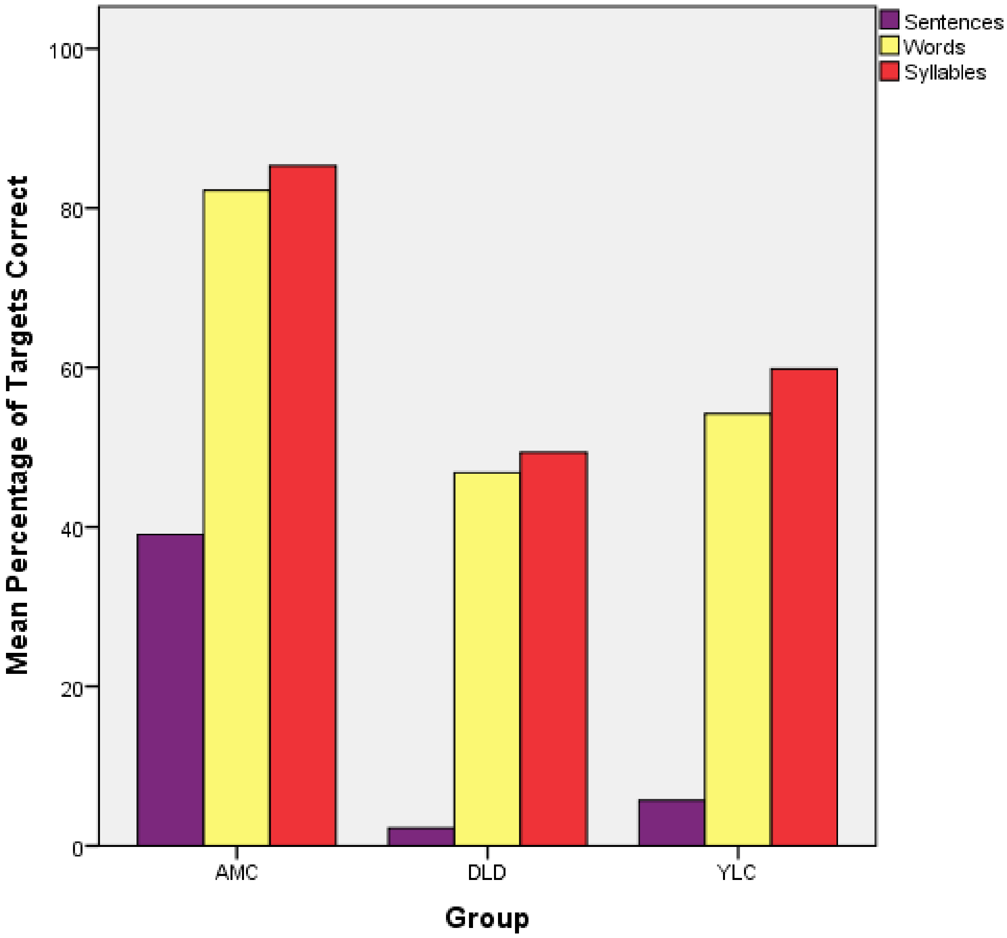


Figure 8-4 Bar Chart of Mean Percentages Correct by linguistic grain-size and participant group

The accuracy scores therefore show that, as expected, the children with DLD had significant difficulties with repetition compared to their AMC peers. They did not differ from the younger, YLC, group at either sentence or word level, which was expected since the DLD and YLC groups were matched partly on the basis of equivalent scores in the Recalling Sentences subtest of the CELF-III^{UK}. At a syllable level of accuracy, however, the DLD group were significantly less accurate than the younger YLC group, indicating that at smaller grain sizes, the difficulties of the DLD group may become more apparent. Any differences in accuracy between the DLD and YLC groups were masked at sentence and word level in this global analysis.

8.8 Results: Effect of Grammatical Status

The next stage of analysis was to investigate whether the grammatical status of the target words and syllables influenced accuracy of repetition.

8.8.1 Effect of Grammatical Status – Content versus Function words

A grammatical hypothesis would predict a significant effect of grammatical status on repetition accuracy, with Content words more accurately repeated than Function words. We therefore split the target words into two groups:

1. Content words were classified as words with the full grammatical and semantic status of nouns, verbs or adjectives (e.g. guitar, hotel, whispered, promised, red, blue).

2. All Function words were classified as determiners (the), auxiliaries (were, had), copulas (is), prepositions (behind, under) and conjunctions (before, after).

8.8.1.1 Effect of Grammatical Status (Content v Function) at Word Level

One-way ANOVAs firstly for Content, and then AllFunction words (DV-Word Score [%Words Correct] in each case) were conducted to see if there were group differences within each category.

Table 8-9 Results of one-way ANOVA for Word Score by Grammatical Status

Word Score (% Words Correct)	**AMC Mean (SD)**	**DLD Mean (SD)**	**YLC Mean (SD)**	***df***	***F***	***p***
Content[a,b]	83.036 (11.892)	45.918 (11.062)	48.292 (16.3085)	2, 58	50.279	.000
All Function[a,b]	81.615 (10.699)	47.411 (12.004)	58.315 (16.279)	2, 58	33.624	.000

Note: a) AMC > DLD b) AMC > YLC

The AMC group were more accurate than the DLD and YLC groups for all measures (*p* = .000). The DLD and YLC groups did not differ in accuracy for Content words, however for AllFunction words the difference between DLD and YLC groups bordered significance (*p*= .057) with the YLC group scoring more highly than the DLD group. Although not significant, this is the first hint of a potential performance difference between the language-impaired children and their younger, language-matched counterparts in repetition skills at the larger grain-size of word-level.

We then conducted a series of t-tests to examine whether accuracy levels differed between Content words and AllFunction words for each participant group.

Table 8-10 Results of t-tests for Word Score by Grammatical Status (Content, AllFunction)

Word Score	**Content**	**AllFunction**	***df***	***t***	***p***
(% Words Correct)	**Mean (SD)**	**Mean (SD)**			
AMC	83.036 (11.892)	81.615 (10.699)	23	.979	.338
DLD	45.918 (11.062)	47.4107 (12.004)	13	-.422	.680
YLC[a]	48.292 (16.308)	58.3152 (16.279)	22	-3.737	.001
All Groups[a]	61.417 (22.074)	64.980 (19.287)	60	-2.312	.024

Note: a) AllFunction > Content

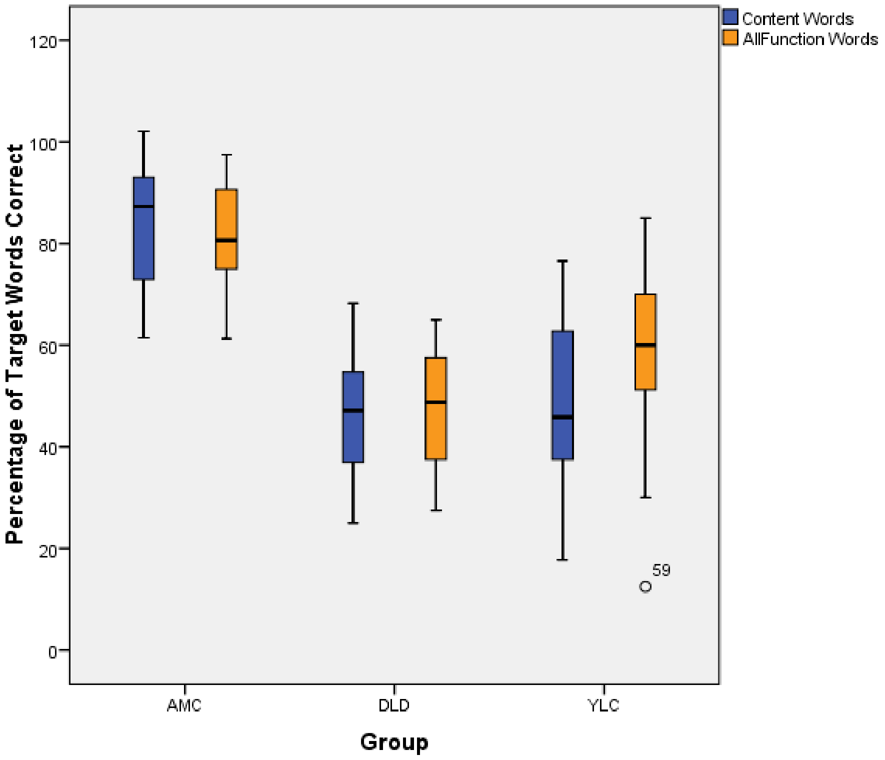


Figure 8-5 Boxplots of Percentage of Target Words Correct by Grammatical Category and Group

We had expected that if there were a difference, it would be that Content words were more accurately repeated than AllFunction words, however there was in fact no significant difference between the grammatical categories for either the AMC or DLD groups, whilst for the YLC group, AllFunction words were significantly more accurate than Content words (*p* = .000), leading to a greater accuracy level for AllFunction words compared to Content words across participants as a whole (*p* = .024).

8.8.1.2 *Effect of Grammatical Status (Content v Function) at Syllable Level*

Having ascertained that grammatical status of Content or Function word did not significantly impact on repetition for AMC and DLD groups at word-level, we were interested to see if a similar pattern emerged at syllable-level. As with the word-level analyses, syllables were therefore assigned to a grammatical group depending on whether they formed part of a Content or Function word.

Table 8-11 Results of one-way ANOVAs for Syllable Scores by Grammatical Status

Syllable Score	**AMC Mean (SD)**	**DLD Mean (SD)**	**YLC Mean (SD)**	***df***	***F***	***p***
Content[a,b]	87.674 (9.175)	51.761 (11.022)	61.156 (16.238)	2,32.128[d]	61.033	.000
All Function[a,b,c]	82.682 (10.515)	47.024 (12.835)	59.239 (17.289)	2, 58	32.873	.000

Note: a) AMC > DLD; b) AMD > YLC; c) YLC > DLD d) Welch's df and F due to significant Levene's test

Post-hoc inspection showed that the AMC group were more accurate than both the DLD and YLC groups in every category (*p* = .000). The YLC group were significantly more accurate than the DLD group in the AllFunction category (*p* = .037). At syllable level, therefore, the accuracy difference between DLD and YLC groups which we saw in the overall syllable score seems to be attributable to poorer accuracy for the DLD group for AllFunction syllables.

Paired samples t-tests were used to explore this further by comparing accuracy rates for Content syllables against AllFunction syllables for each participant group.

Table 8-12 Results of t-tests for Syllable Score by Grammatical Category (Content, AllFunction)

Syllable Score	Content Mean (SD)	AllFunction Mean (SD)	*df*	*t*	*p*
AMC[a]	87.674 (9.175)	82.682 (10.515)	23	4.862	.000
DLD	51.761 (11.022)	47.024 (12.835)	13	1.336	.220
YLC	61.156 (16.238)	59.239 (17.289)	22	.793	.436

Note: a) Content > AllFunction

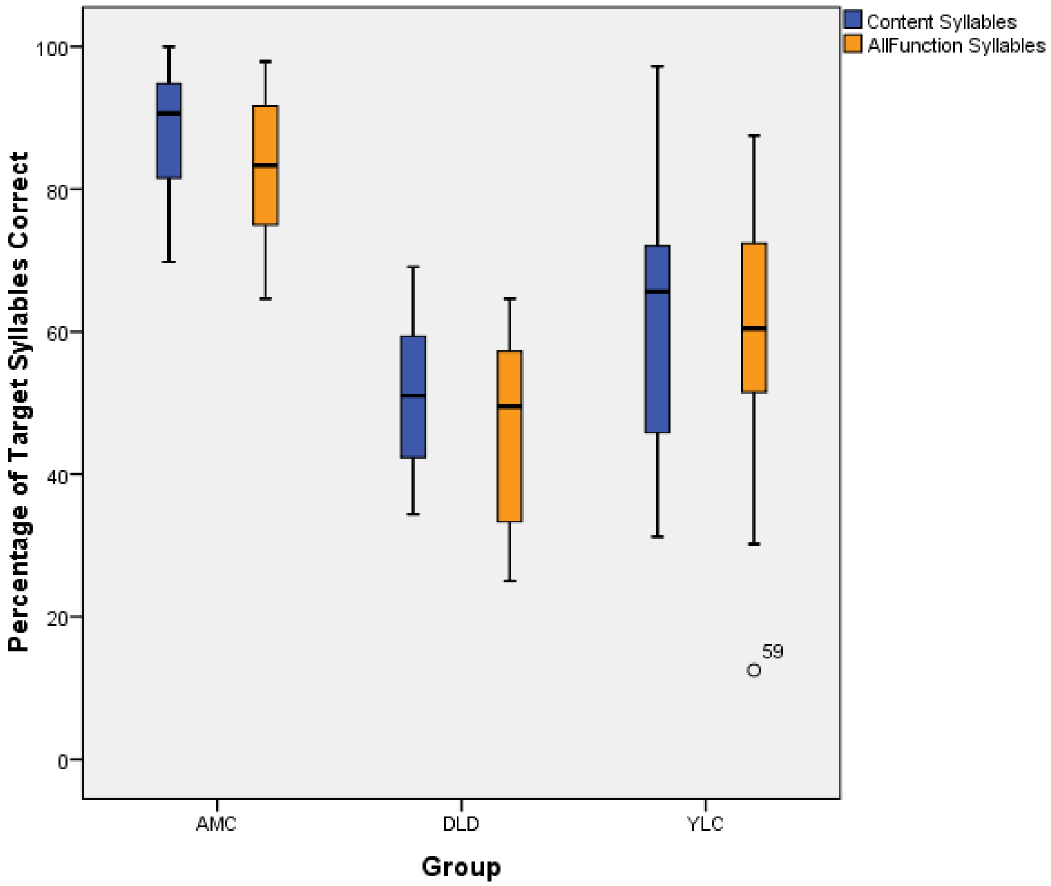


Figure 8-6 Boxplots of Syllable Scores for Content and Function Syllables by Group

The AMC group were significantly more accurate at repeating Content than AllFunction syllables (*p* = .000), however the other two groups did not significantly differ in accuracy between the grammatical categories.

8.8.1.3 Effect of Grammatical Status (Content v Function) Conclusions

Previous accounts of sentence repetition have noted a significant effect of grammatical status of target words on repetition accuracy, with Content words more accurately repeated than Function words. These were not, however, the results that we obtained. At word level, there were no significant differences in repetition between Content words and Function words for AMC or DLD groups. The YLC group did show a significant difference, but in the opposite direction than expected. Their repetition of Function words was more accurate than that of Content words.

Previous work has not investigated sentence repetition at the syllable-level, however it seemed reasonable to predict a similar effect: that syllables in Content words would be more accurately repeated than syllables in Function words. The AMC group displayed the predicted effect with Content-word syllables more accurate than Function-word syllables, however there was no effect of grammatical status for either the DLD or YLC groups.

The emerging picture is therefore more complex than a straightforward grammatical theory would have predicted. At word-level, YLC were the only group to show a grammatical effect, with Function words more accurate than Content words. In contrast, at syllable-level, AMC were the only group to show a grammatical effect, with Content syllables more accurate than Function syllables. The DLD group showed no effect of grammatical status at either level of analysis.

8.8.2 Effect of Grammatical Status - Content Words versus Subdivided Function Word Categories

Previous work has categorised grammatical status by dividing words into either Content or Function words. Because of our interest in stress and prosody however, we wondered whether this categorisation may be obscuring intra-categorical prosodic distinctions, particularly with regard to the Function category. Determiners and auxiliaries are typically unstressed, whereas Function words such as prepositions and conjunctions may contain both stressed and unstressed syllables. To investigate whether Function words containing stressed syllables behaved differently to their unstressed counterparts, we divided the AllFunction category into two subcategories:

1. ‘**Detaux’ Function words** consisted of **det**erminers, **aux**iliaries and copulas – all unstressed syllables

2. **PrepConj words** consisted of **prep**ositions (behind, under) and **conj**unctions (before, after) – all containing a stressed syllable

8.8.2.1 Effect of Grammatical Status - Content versus Detaux and PrepConj – Word Level

The analysis began at the larger grain-size of word-level. A series of one-way ANOVAs (DV – Word Score [% Words correct]) was conducted for each grammatical category to determine any differences between participant groups.

Table 8-13 Results of one-way ANOVAs for each grammatical category at word-level

Word Score (% Words Correct)	**AMC Mean (SD)**	**DLD Mean (SD)**	**YLC Mean (SD)**	***df***	***F***	***p***
Content[a,b]	83.036 (11.892)	45.918 (11.062)	48.292 (16.309)	2, 58	50.279	.000
Detaux[a,b]	80.599 (10.873)	49.554 (11.407)	57.948 (15.367)	2, 58	31.167	.000
PrepConj[a,b,c]	85.677 (12.565)	38.839 (20.093)	59.783 (23.296)	2, 29.463[d]	35.280	.000

Note: a) AMC > DLD; b) AMD > YLC; c) YLC > DLD d) Welch’s df and F due to significant Levene’s test

Post-hoc inspection showed that the AMC group were more accurate than the DLD and YLC groups for all measures (*p* = .000). We had previously noted that the difference between DLD and YLC groups bordered significance for AllFunction words. Once the AllFunction category was subdivided, we found that the DLD group were not significantly different from the YLC group for Detaux words, but were significantly less accurate than the YLC group for PrepConj words (*p* = .019), which suggests that this latter difference is what drove the AllFunction result.

To investigate whether Content, Detaux and PrepConj words differed in accuracy for each group, a series of one-way repeated-measures ANOVAs on Grammatical Status [Content, Detaux, PrepConj] – DV % Words Correct) were conducted.

Table 8-14 Results of repeated-measures ANOVAs by Grammatical Status

Word Score	**Content Mean (SD)**	**Detaux Mean (SD)**	**PrepConj Mean (SD)**	***df***	***F***	***p***
AMC[a]	83.036 (11.892)	80.599 (10.873)	85.677 (12.565)	2, 46	4.152	.022
DLD	45.918 (11.062)	49.554 (11.407)	38.839 (20.093)	2, 26	3.209	.057
YLC[b,c]	48.292 (16.308)	57.948 (15.367)	59.783 (23.296)	1.481, 32.592[d]	7.073	.006
All Groups	61.417 (22.074)	64.9334 (18.225)	65.164 (26.232)	1.643, 95.309 [d]	2.133	.123

Note: a) PrepConj > Detaux; b) Detaux > Content c) PrepConj > Content d) Greenhouse-Geisser correction due to significant Mauchly's test

Both the AMC and YLC group showed a significant effect of grammatical status. Pairwise comparisons (Bonferroni) revealed that the AMC group repeated PrepConj words significantly more accurately than Detaux words (*p* = .024), whereas for the YLC group Detaux and PrepConj words were significantly more accurate than Content words (*p* = .003 and *p* = .032 respectively). The DLD group scores did not significantly differ by grammatical category (*p* = .057).

A repeated-measures ANOVA (3 x 3 – Group [AMC,DLD,YLC] x Grammatical status [Content,Detaux,PrepConj] confirmed differing group response patterns with a significant Group*Grammatical status interaction *F*(4) = 6.597, *p* = .000.

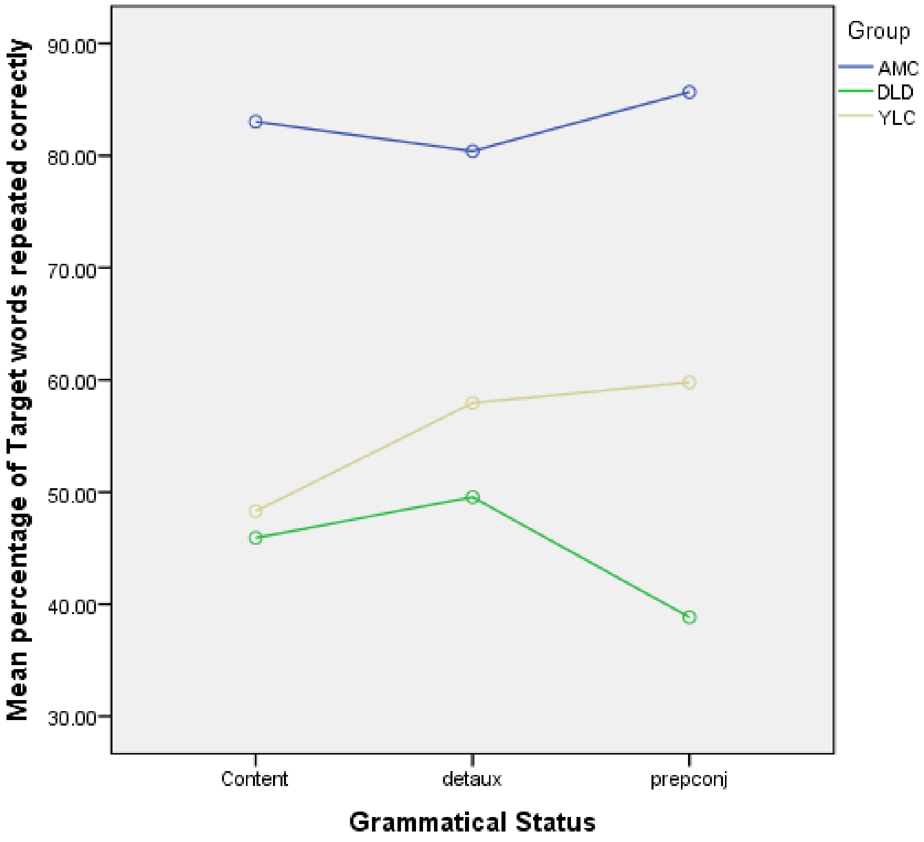


Figure 8-7 Graph showing mean percentage of words correct by group and grammatical status

Figure 8-7 illustrates the different response patterns. The DLD group have a slight increase in accuracy between Content and Detaux words with decreased accuracy for PrepConj words, whilst for the AMC group, PrepConj are their most successful repetition category. For the YLC group it was Content words that were least accurate.

We had expected that Content words would be better repeated than Detaux, however this was not the case, with no significant results in that direction, and in fact the opposite was true of the YLC group, who repeated Detaux words significantly more accurately than Content words. AMC was the only group to show a significant difference within the AllFunction category, being more accurate for PrepConj words than for Detaux words.

We had hypothesised that Content words would be more accurate than Function words and that within Function words, PrepConj would be more accurate than Detaux due to the presence of a stressed syllable. It can be seen from the graph that this was broadly the pattern of results for the AMC group, although only the Detaux:PrepConj difference was significant. For the YLC group, however, both Detaux and PrepConj words were more accurate than Content words. For the DLD group, despite a visible trend towards poorer PrepConj repetition, there was no significant effect of grammatical status.

8.8.2.2 *Effect of Grammatical Status – Content versus Detaux and PrepConj Syllable Level*

To continue the investigation of the two sub-categories of Function Word, we examined these categories again at syllable level by conducting a series of one-way repeated-measures ANOVAs for Grammatical Status [Content, Detaux, PrepConj] – DV % Syllables Correct).

Table 8-15 Results of repeated-measures ANOVAs for Syllable Score by Grammatical Status (Content, Detaux, PrepConj)

Syllable Score (% Correct)	**Content**	**Detaux**	**PrepConj**	***df***	***F***	***p***
AMC[a,b]	87.674 (9.174)	80.599 (10.874)	86.849 (12.294)	1.542, 35.463[c]	11.533	.000
DLD	51.761 (11.022)	49.554 (11.407)	41.964 (20.390)	2, 26	2.769	.081
YLC	61.156 (16.238)	57.948 (15.367)	61.821 (24.058)	1.531, 33.675[c]	.876	.423
AllGroups[a]	68.989 (19.886)	64.933 (18.225)	67.111 (25.968)	1.636, 94.911[c]	3.277	.052

Note: a) Content > Detaux; b) PrepConj > Detaux; c) Greenhouse-Geisser correction due to significant Mauchly's test.

There was a significant effect of Grammatical Status for the AMC group, with Content syllables repeated significantly more accurately than Detaux (p = .000) and PrepConj syllables also repeated significantly more accurately than Detaux syllables (*p* = .008) (Bonferroni comparisons), driving the AllGroups effect towards significance. There were no significant effects of grammatical status for the DLD or YLC groups.

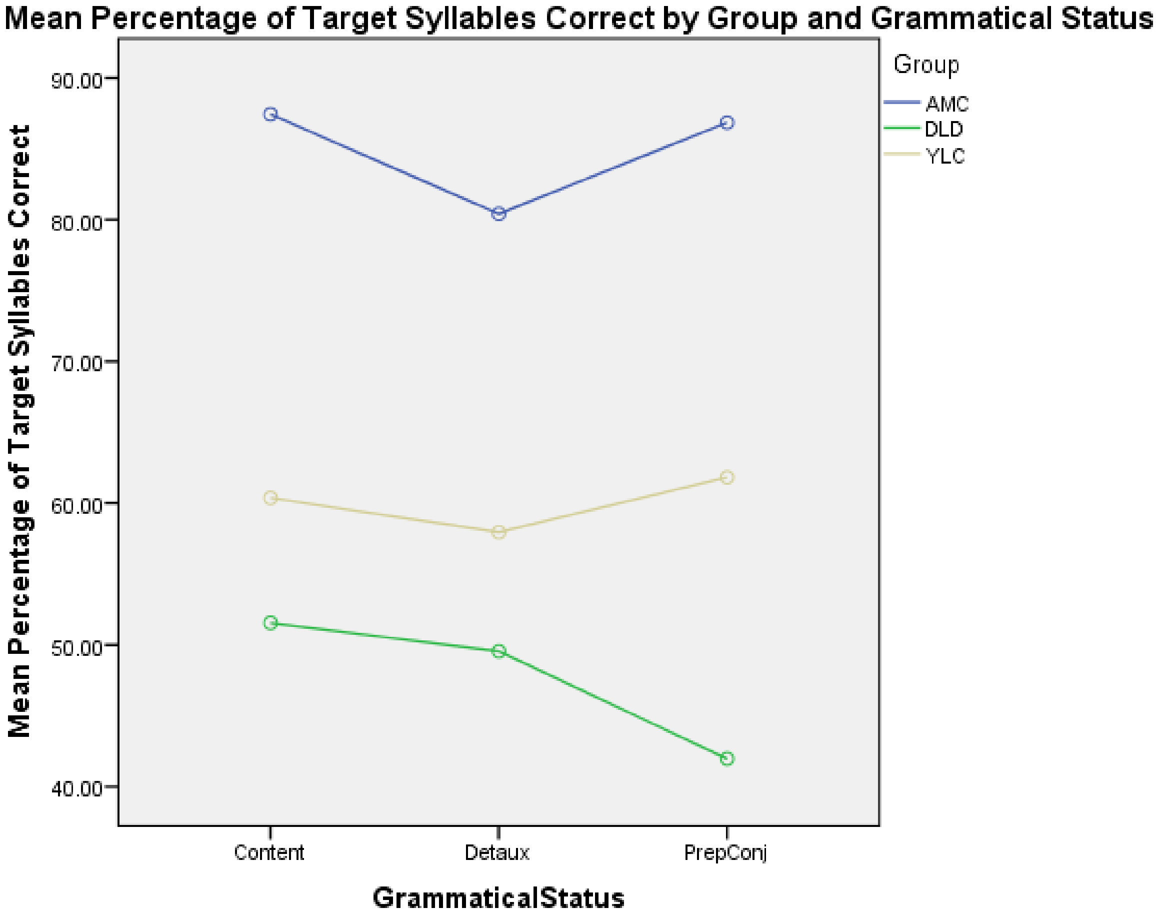


Figure 8-8 Graph showing mean percentage of target syllables correct by group and grammatical status

The differing response patterns for the three participant groups were confirmed by a repeated-measures ANOVA (3 x 3 – Group [AMC, DLD, YLC] x Grammatical Status [Content, Detaux, PrepConj]) which showed a significant Group* Grammatical Status Interaction *F*(4) = 3.052, *p* = .02. Figure 8-8 shows that although the YLC group showed no significant effect of Grammatical Status, their response pattern nonetheless appears to follows a similar trend to that of the AMC group. In contrast, the pattern for the DLD group differs considerably, with a drop, rather than increase, in accuracy for the PrepConj syllables.

These syllable-level results mirror the results of the word-level analysis, in that grammatical status is a significant factor in the AMC group's repetition accuracy, but not for the DLD group. For the YLC group, their lower accuracy for Content words was not replicated for Content syllables, with no effect of grammatical status at syllable-level for this group.

We had expected that Content syllables would be more accurate than Detaux syllables, as was found for the AMC group. Interestingly, the PrepConj syllables (which, it will be remembered, also included stressed syllables) were also more accurate than Detaux syllables for this group despite both PrepConj and Detaux forming part of the overall Function word category. This hints that rhythmic properties may be a more significant factor in repetition for this group than grammatical category. This will be examined in more detail in Section 8.9.3

8.8.3 Summary of Effect of Grammatical Status on Accuracy

A grammatical theory would predict that Content words would be more accurately repeated than AllFunction words. This result did not obtain for any of our participant groups, however at syllable-level the AMC group had significantly more accurate repetition for Content than AllFunction syllables, following the predicted direction of scores at a smaller grain-size than sentence or word level. In contrast, the YLC group showed a significant word-level result in the opposite direction, with AllFunction words being more accurate than Content words.

By dividing AllFunction items into two sub-categories, we were able to look at grammatical status in more detail. This produced a more complex result, with the AMC group displaying a significant difference in accuracy between PrepConj and Detaux words and syllables – i.e. a differentiation of scores within the AllFunction category. Furthermore, by looking across scores, the significant difference between Content and AllFunction words for this group appears to have been driven by a particular vulnerability for Detaux words and syllables, rather than lesser accuracy across the AllFunction category as a whole. This indicates that a broad division into Content and Function may mask some interesting intra-category distinctions.

The YLC group did not differ across grammatical categories at syllable level, but at word level had significantly better AllFunction scores, which proved to be based on better accuracy for both Detaux and PrepConj words in comparison to Content words. The relatively higher accuracy at syllable-level for the Content category suggests a syllable-level factor supporting repetition. Such a factor could be the presence of a stressed syllable supporting repetition of individual syllables but not operating at word level.

The most striking outcome is that there was no effect of grammatical status for the DLD group across word and syllable levels, regardless of whether the comparison was Content against AllFunction or against the sub-divided Detaux and PrepConj categories. A grammatical hypothesis would predict that certain grammatical categories (i.e. AllFunction or Detaux) would be particularly vulnerable for this group, however the non-significant trend was towards a particular difficulty with the PrepConj category. This was not a significant difference, however, so we must conclude that grammatical status did not affect repetition accuracy for the language-disordered group.

8.9 Results: Effect of Stress Status

The second hypothesis we wanted to investigate was that the stress status of a syllable as strong or weak would impact on repetition accuracy, with stressed syllables more accurate than unstressed syllables. To do this, we divided all syllables into two groups – stressed (those carrying primary stress e.g. TEL of hoTEL) and unstressed (e.g. ho of hoTEL). The score was calculated as the number of syllables repeated correctly expressed as a percentage of the total number of target syllables (maximum score therefore 100).

8.9.1 Effect of Stress on Syllable-Level Accuracy

Table 8-16 Results of t-tests for Stressed and Unstressed Syllable Scores by Group

Syllable Score (% Correct)	**Stressed**	**Unstressed**	***df***	***t***	***p***
AMC[a]	88.831 (8.17563)	83.5069 (10.23749)	23	6.670	.000
DLD[a]	52.9762 (11.09584)	48.2639 (10.15063)	13	2.849	.014
YLC	55.5556 (17.05153)	57.8200 (15.500)	22	-1.721	.099
AllGroups[a]	68.0556 (21.08338)	65.7332 (19.29168)	60	2.807	.007

Note: a) Stressed > Unstressed

A series of paired samples t-tests showed that stressed syllables were more accurately repeated than unstressed by the AMC (p = .000) and DLD groups, (p = .014), whilst the YLC group showed no effect of stress status (p = .099). A repeated-measures ANOVA (3 x 2 – Group [AMC, DLD, YLC] x Stress [stressed, unstressed] confirmed the differing pattern for the YLC group, with a significant stress*group interaction ($F(2)$ = 13.111, p = .000) (see Figure 8-9).

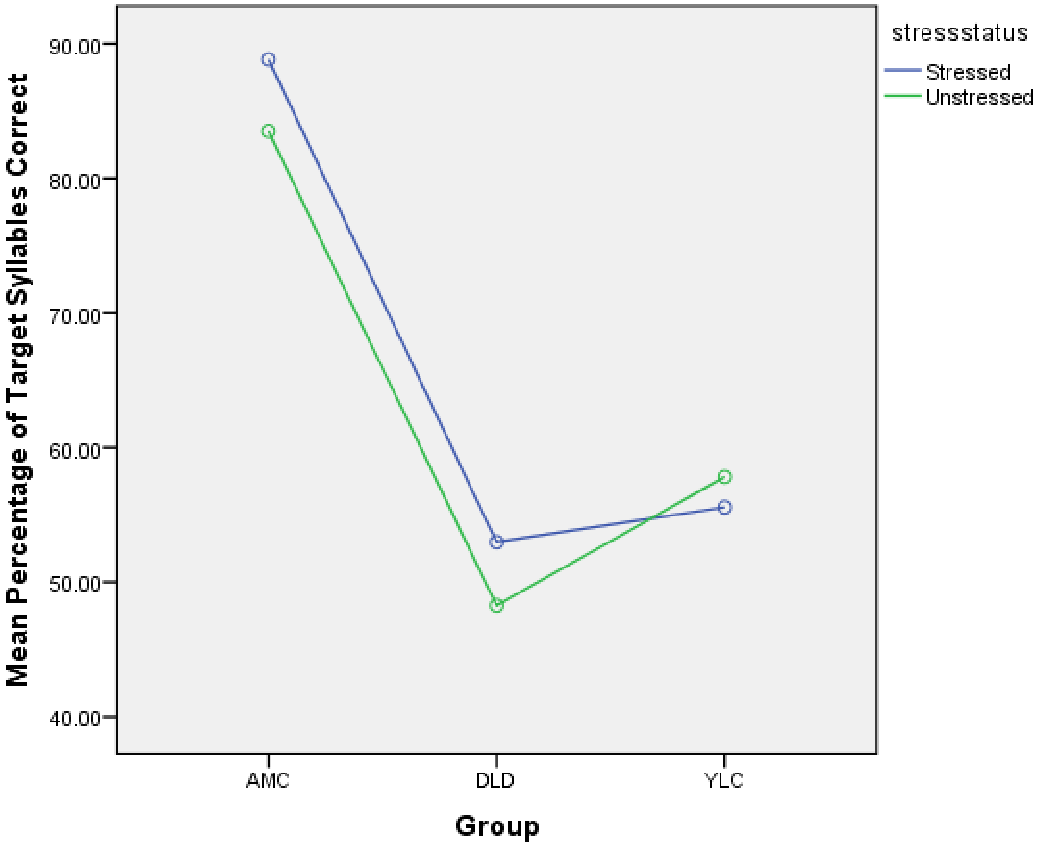


Figure 8-9 Graph showing Mean Percentage of Target Syllables Correct by Stress Status

Stress status therefore has a significant impact on repetition at the syllable level, with stressed syllables being consistently more accurately repeated than unstressed syllables by both AMC and DLD groups. In contrast, accuracy for the YLC group did not differ between stressed and unstressed syllables.

8.9.2 The Effect of Stress and Grammatical Status on Syllable Level Accuracy

Following the overall Stress analysis, which showed a significant effect of stress on syllable-level accuracy for the AMC and DLD groups, we were interested to know if this altered with the grammatical status of the target syllable (i.e. whether it was part of a content word or a function word). We therefore allocated the syllables as either stressed or unstressed (stress status) and as Content, Detaux or PrepConj (grammatical status). For example, ‘TEL’ of ‘hoTEL’ would be ‘stressed, Content’, whereas ‘be’ of ‘beFORE’ would be ‘unstressed, PrepConj’. It should be remembered that all Detaux syllables were unstressed.

8.9.2.1 *Comparing Grammatical Status across Stressed Syllables*

Firstly, Stressed Syllable Scores were analysed to investigate any impact of grammatical status on stressed syllable repetition accuracy. Only Content and PrepConj words contained a stressed syllable (all Detaux syllables were unstressed), so only these two categories were included in the analysis.

Table 8-17 Results of t-tests comparing Stressed Syllable Scores for Content and PrepConj words

Stressed Syllable Score (% Correct)	**Content**	**PrepConj**	***df***	***t***	***p***
AMC	88.914 (8.523)	87.5 (12.5)	23	.742	.466
DLD	53.954 (10.322)	43.304 (21.576)	13	1.967	.071
YLC	64.208 (15.949)	63.044 (24.564)	22	.310	.759
All Groups	71. 575 (18.913)	68.135 (26.188)	60	1.676	.099

A series of t-tests were conducted comparing Content and PrepConj syllables for each group (DV - % Stressed Syllables correct). There were no significant effects of grammatical status for any of the groups. Accuracy of stressed syllables was therefore not affected by whether the target syllable occurred as part of a Content or a PrepConj word, suggesting that grammatical category was not a significant factor in repetition accuracy for stressed syllables. However, a trend is there for the DLD group (*p* = .071).

8.9.2.2 Effect of Grammatical Status across Unstressed syllables

If grammatical status (Content, Function) had a significant influence on repetition, then we would expect that unstressed syllables in both Detaux and PrepConj categories would be less accurately repeated than unstressed syllables in Content words.

To investigate this, we conducted a one-way repeated-measures ANOVA for Grammatical Status [Content, Detaux, PrepConj], DV - % Unstressed Syllables Correct)

Table 8-18 Results of one-way repeated-measures ANOVA for Unstressed Syllable Score (% Correct) for Content, Detaux and PrepConj words

Unstressed Syllable Score (% Correct)	**Content Mean (SD)**	**Detaux Mean (SD)**	**PrepConj Mean (SD)**	***df***	***F***	***p***
AMC[a,b]	85.938 (10.677)	80.599 (10.873)	86.198 (12.361)	2, 46	7.799	.001
DLD	48.772 (12.214)	49.554 (11.407)	40.625 (19.878)	2, 26	2.812	.078
YLC	56.929 (17.525)	57.948 (15.367)	60.598 (23.794)	2, 44	.655	.524
All Groups	66.470 (21.171)	64.933 (18.225)	66.0861 (26.007)	1.8, 104.410	.376	.666

Note: a) Content > Detaux; b) PrepConj > Detaux

There was an effect of grammatical status for the AMC group only, with unstressed Content syllables and unstressed PrepConj syllables more accurate than unstressed Detaux syllables (*p* = .001, *p* = .014 respectively, Bonferroni comparisons). In contrast there was no significant effect of grammatical status for either DLD or YLC groups.

A repeated-measures ANOVA (3 x 3 – Group [AMC, DLD, YLC] x Grammatical Status [Content, Detaux, PrepConj] confirmed that the response patterns differed between the groups with significant group*grammatical status interactions: *F*(4) = 3.586, *p* = .009.

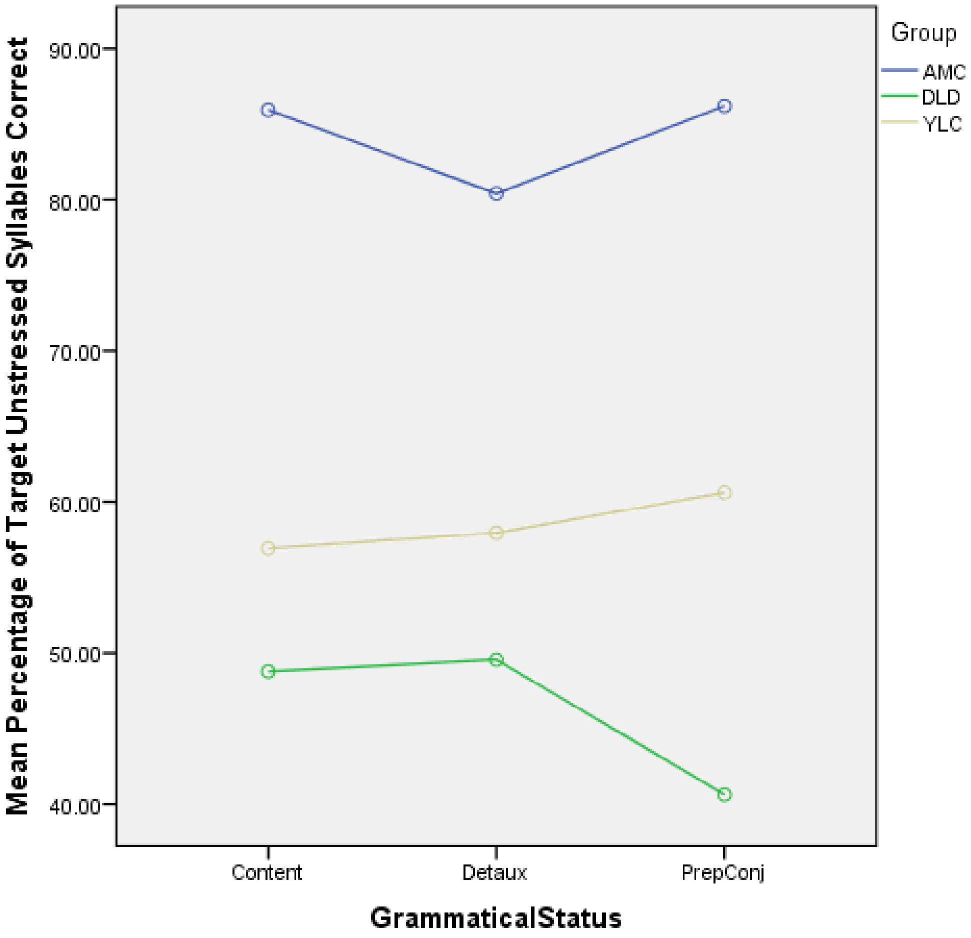


Figure 8-10 Graph showing mean percentage of target unstressed syllables correct by grammatical status and group

Unstressed syllables in Content and PrepConj words were therefore more accurately repeated by the AMC group than unstressed syllables in Detaux words. This indicates that a binary division into Content and Function does not adequately capture the data, since PrepConj and Detaux both fall into the Function category. Conversely, both PrepConj and Content words contain a Stressed syllable. The data seem to point towards a rhythmic interpretation, in which unstressed syllables occurring within words which also carry a stressed syllable (Content and PrepConj) are better preserved than unstressed syllables occurring in isolation (Detaux) for typically-developing children.

8.9.3 Effect of Stress on Word Level Accuracy

Following this finding, we hypothesised that the presence of a stressed syllable within a word might facilitate whole-word repetition. We therefore created two prosodically-motivated categories: Words with a stressed syllable (i.e. Content and Prepconj) and Words without a stressed syllable (i.e. Detaux) by taking the mean percentage score across the component variables. We then conducted a series of paired sample t-tests to see if the two categories differed in word-level accuracy.

Table 8-19 Results of paired sample t-tests for words with and without stress

Word Score	**Words with a Stressed syllable Mean (SD)**	**Words without a Stressed syllable Mean (SD)**	***df***	***t***	***p***
AMC[a]	84.356 (11.199)	80.599 (10.873)	23	2943	.007
DLD[b]	42.379 (13.367)	49.554 (11.407)	13	-2.306	.038
YLC[b]	54.037 (17.508)	57.948 (15.367)	22	-2.119	.046
All Groups	63.290 (22.617)	64.857 (18.138)	60	-1.324	.190

Note: a) WordsWithStress > WordsWithoutStress; b)WordsWithoutStress > WordsWithStress

For the AMC Group, words containing a stressed syllable were more accurately repeated than those without a stressed syllable (*p* = .004), reflecting the greater accuracy for this group in repeating both Content and PrepConj words than Detaux words. The DLD and YLC groups, however, displayed the opposite finding, as for both these groups, words without a stressed syllable (i.e Detaux) were more accurate than words with a stressed syllable.

The presence of a stressed syllable within a word may therefore contribute to the repetition accuracy for the AMC group, but it does not have a beneficial effect on repetition for the DLD and YLC groups, for whom words without a primary stress were more accurately repeated. The beneficial effect of stress for the DLD group that we saw at syllable-level for individual syllables carrying stress therefore does not translate to a broader word-level advantage for entire words containing primary stress.

8.9.4 Summary of effect of Stress on Accuracy

We predicted that stressed syllables would be more accurately repeated than unstressed syllables. This was the result we obtained for AMC and DLD groups, however there was no difference between stressed and unstressed syllable accuracy for the YLC group. The advantage for stressed over unstressed syllables did not vary according to grammatical category, suggesting that prosodic influence on accuracy applies independently of grammatical status.

The results of the AMC group reveal a word-level effect of stress, as words containing a stressed syllable were more accurately repeated than those without a stressed syllable. Since stressed syllable accuracy did not differ between grammatical categories, this suggests that the difference in accuracy for unstressed syllables was driving this effect. It seems that for typically-developing children, the presence of a stressed syllable in a word has a facilitating effect on repetition of that whole word, leaving words comprised of a single unstressed syllable more vulnerable.

The accuracy advantage for stressed syllables did not translate to a greater accuracy for whole-words containing stress for the DLD group. This may reflect a difficulty in integrating individual stress levels into a larger rhythmic structure operating at word level. There may also be an interaction between rhythmic and grammatical influences, since the stress-containing PrepConj category reflects a relatively advanced grammatical structure and proved to be particularly problematic for the DLD children.

8.10 Effect of Entrainment

As will be recalled, the AMC and DLD groups also completed an Entrained version of the task. We predicted that listening to an entraining beat would provide a beneficial scaffold for the rhythmic structure of the stimulus sentence and thus facilitate repetition. We therefore predicted that scores for Entrained sentences would be higher than those for Unentrained sentences.

8.10.1 Effect of Entrainment on overall Accuracy

Scores for the Entrained sentences were calculated as for the Unentrained versions for accuracy at sentence, word and syllable level.

Table 8-20 Sentence, Word and Syllable Scores for Unentrained and Entrained versions

Score (% Target Correct)	**AMC Mean (SD)**		**DLD Mean (SD)**	
	Unentrained	**Entrained**	**Unentrained**	**Entrained**
Sentence Score	39.063 (22.364)	41.146 (24.165)	2.232 (5.262)	2.232 (3.108)
Word Score	82.2 (10.644)	83.578 (9.980)	46.796 (9.633)	46.639 (12.421)
Syllable Score	85.262 (9.633)	86.651 (9.938)	49.339 (9.950)	49.173 (12.451)

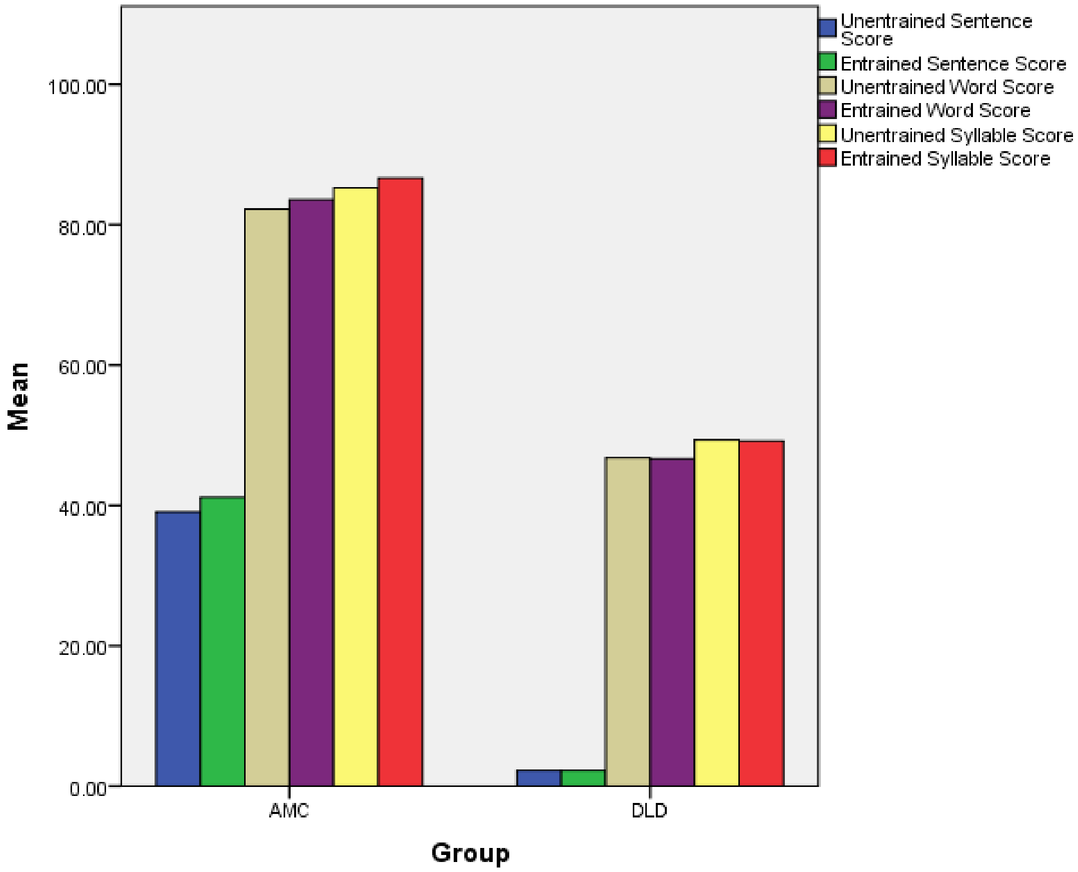


Figure 8-11 Graph of the Mean Sentence, Word and Syllable Level Scores for both Unentrained and Entrained versions

A repeated-measures ANOVA was conducted for each grain-size (2 x 2 – Group [AMC, DLD] x Entrainment [Unentrained, Entrained]; DV-Score). There was no significant main effect of entrainment at Sentence Level, *F* (1, 36) = .018, *p* = .894, Word Level, *F*(1, 36) = .177, *p* = .676, or at Syllable Level *F*(1, 36) = .305, *p* = .584.

Contrary to our predictions therefore, the entraining beat did not facilitate more accurate repetitions at any of the levels of analysis.

8.10.2 Effect of Entrainment with regard to Grammatical Status

As part of our analysis we also looked at whether Entrainment had differing effects according to grammatical category.

8.10.2.1 Effect of Entrainment with regard to Grammatical Status (Content v AllFunction)

We began by examining the Content and AllFunction categories, first at word- and then at syllable-level, using a repeated-measures ANOVA (2 x 2 x 2 – Group [AMC, DLD] x Grammatical Status [Content, AllFunction] x Entrainment [Entrained, Unentrained] – DV %WordsCorrect; % SyllablesCorrect respectively. A 2 x 2 ANOVA was then conducted for each group individually.

Table 8-21 Results of repeated-measures ANOVAs - Grammatical Status (Content, AllFunction) x Entrainment – Word Level

Word Level (Content, AllFunction) - % Words Correct		***df***	***F***	***p***
All Groups				
Main Effects	**Group**	1, 36	117.767	.000
	Entrainment	1, 36	.169	.683
Interaction	**Grammatical Status*Entrainment**	1, 36	2.043	.162
AMC only				
Main Effect	**Entrainment**	1, 23	.585	.452
Interaction	**Grammatical Status*Entrainment**	1, 23	.126	.726
DLD only				
Main Effect	**Entrainment**	1, 13	.064	.805
Interaction	**Grammatical Status*Entrainment**	1, 13	.324	.579

Syllable Level (Content, AllFunction) - % Syllables Correct		***df***	***F***	***p***
All Groups				
Main Effects	**Group**	1, 36	120.019	.000
	Entrainment	1, 36	.751	.392
Interaction	**Grammatical Status*Entrainment**	1, 36	.062	.804
AMC only				
Main Effect	**Entrainment**	1, 23	.696	.413
Interaction	**Grammatical Status*Entrainment**	1, 23	.000	.983
DLD only				
Main Effect	**Entrainment**	1, 13	.247	.627
Interaction	**Grammatical Status*Entrainment**	1, 13	.067	.800

There were no significant effects of Entrainment for any of the groups and nor did Entrainment have any significant interactions with Grammatical Status (Content, AllFunction). Repetition accuracy of different grammatical categories was therefore not differentially affected by listening to an entraining beat at either word or syllable level.

8.10.2.2 Effect of Entrainment with regard to Grammatical Status – Content v Detaux/PrepConj

We then looked at whether Entrainment had a significant effect when taking the sub-categories of Detaux and PrepConj into account. Again, we looked at both word- and syllable-level scores, using a repeated measures ANOVA (2 x 3 x 2 – Group [AMC,DLD] x Grammatical Status [Content, Detaux, PrepConj] x Entrainment [Unentrained, Entrained], then repeating the ANOVA as a 3 x 2 for each group individually (Table 16).

As with the Content/AllFunction results, there were no significant effects of Entrainment and no significant interactions of Entrainment with Grammatical Status (Content, Detaux, PrepConj) at either word or syllable level.

Listening to an entraining beat did not therefore significantly affect accuracy and nor did entrainment status differ according to the grammatical status of target words.

Table 8-22 Results of repeated-measures ANOVAs - Grammatical Status (Content, Detaux, PrepConj) x Entrainment – Word & Syllable Level

Word Level (Content, Detaux, PrepConj) - % Words Correct		***df***	***F***	***p***
All Groups				
Main Effects	**Group**	1, 36	110.661	.000
	Entrainment	1, 36	.233	.632
Interaction	**Grammatical Status*Entrainment**	2, 72	.238	.788
AMC only				
Main Effect	**Entrainment**	1, 23	.294	.593
Interaction	**Grammatical Status*Entrainment**	2, 46	.764	.472
DLD only				
Main Effect	**Entrainment**	1, 13	.037	.851
Interaction	**Grammatical Status*Entrainment**	2, 26	.347	.710

Syllable Level (Content, Detaux, PrepConj) - % Syllables Correct		***df***	***F***	***p***
All Groups				
Main Effects	**Group**	1, 36	109.383	.000
	Entrainment	1, 36	.695	.410
Interaction	**Grammatical Status*Entrainment**	2, 72	.035	.965
AMC only				
Main Effect	**Entrainment**	1, 23	.349	.561
Interaction	**Grammatical Status*Entrainment**	2, 46	1.670	.200
DLD only				
Main Effect	**Entrainment**	1, 13	.435	.521
Interaction	**Grammatical Status*Entrainment**	2, 26	.356	.704

8.10.3 Effect of Entrainment with regard to Stress status at Syllable Level

In order to see whether Entrainment had differing impacts according to whether syllables were stressed or unstressed, we conducted a repeated-measures ANOVA (2 x 2 x 2 – Group [AMC,DLD] x Stress Status [stressed,unstressed] x Entrainment [Unentrained, Entrained] – DV % Syllables Correct), then repeated as 2 x 2 ANOVAs for each group individually.

Table 8-23 Results of repeated-measures ANVOVAs – Stress Status (Stressed, Unstressed) x Entrainment

Syllable Level (Stressed, Unstressed) - % Syllables Correct		***df***	***F***	***p***
All Groups				
Main Effects	**Group**	1, 36	128.328	.000
	Entrainment	1, 36	.192	.664
Interaction	**Stress Status*Entrainment**	1, 36	.002	.968
AMC only				
Main Effect	**Entrainment**	1, 23	.656	.426
Interaction	**Stress Status*Entrainment**	1, 23	.715	.407
DLD only				
Main Effect	**Entrainment**	1, 13	.031	.863
Interaction	**Stress Status*Entrainment**	1, 13	.142	.713

There were no significant effects of Entrainment and nor did Entrainment interact with Stress Status in any of the results. Listening to an entraining beat therefore did not differentially impact on accuracy for either stressed or unstressed syllables.

8.10.4 Summary of Effect of Entrainment

We had hypothesised that providing an entraining beat would facilitate accuracy of repetition, however this was not the case in our results. There was no significant effect of Entrainment on accuracy at sentence, word or syllable levels for the participants as a whole or for individual groups. Nor was there a significant interaction of Entrainment status with either of the other variables of interest i.e. grammatical and stress status. We must therefore conclude that listening to an entraining beat, at least of the type used here, does not affect accuracy of sentence repetition.

8.11 Effect of Immediate Prosodic Context on Function Word Accuracy

The third hypothesis we wanted to explore was that investigated by Gerken (1996) and McGregor & Leonard (1994). They hypothesised that children were more likely to retain unstressed Function words if they occurred in a Sw context – i.e. if the target Function word was preceded by a strong syllable. We designed our stimuli to test this hypothesis by creating sentences in which the unstressed Function words systematically occurred in either a Sw context or in a ww context.

Table 8-24 Example Sentences illustrating the location of unstressed Function words – Embedded Sentences

Row	**Type**	**Example Sentence**												**Rhythm**
1	**Emb-Reg-ww**	the	**CAR**	pet	the	**ME**	tal	is	**UN**	der	is	**PUR**	ple	regular
2	**Emb-Var-ww**	the	**TA**	bles	the	com	**PU**	ter	is	**UN**	der	are	**BIG**	variable
3	**Emb-Reg-Sw**	the	dis	**PLAY**	the	gui	**TAR**	is	be	**HIND**	is	com	**PLETE**	regular
4	**Emb-Var-Sw**	the	ma	**CHINE**	the	ma	ga	**ZINE**	is	be	**HIND**	is	**SHORT**	variable

Note: purple shading = Function word following unstressed syllable i.e. ww context; yellow shading = Function word following stressed syllable i.e. Sw context

Table 8-25 Example Sentences illustrating the location of unstressed Function words – Complex Sentences

Type	**Example Sentences**														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Com-Reg-ww	the	**STU**	dents	were	**AR**	gu	ing	**AF**	ter	the	**MEE**	ting	had	**STAR**	ted
Com-Var-ww	the	ex	**PLO**	rer	was	**RES**	ting	**AF**	ter	the	ad	**VEN**	ture	had	**STOPPED**
Com-Reg-Sw	the	ho	**TEL**	was	co	**LLAP**	sing	be	**FORE**	the	po	**LICE**	had	pre	**PARED**
Com-Var-Sw	the	**TEAM**	were	**CE**	le	bra	ting	be	**FORE**	the	**MATCH**	had	con	**CLU**	ded

Note: purple shading = Function word following unstressed syllable i.e. ww context; yellow shading = Function word following stressed syllable i.e. Sw context

We then created two variables – wwContext and SwContext – by taking the mean percentage score for each of the location types and used paired sample t-tests to look for accuracy differences. Since entrainment had no significant impact on scores (see Section 8.10), the percentage correct from both unentrained and entrained scores was taken[23].

---

[23] The YLC group's scores are percentages based on unentrained scores only.

Table 8-26 Results of paired sample t-tests for wwContext and SwContext

Score	**wwContext** **Mean (SD)**	**SwContext** **Mean (SD)**	***df***	***t***	***p***
AMC[a]	85.677 (12.215)	79.861 (12.615)	23	3.865	001
DLD[a]	48.81 (14.288)	39.583 (15.287)	13	2.706	.018
YLC[a]	64.402 (17.4)	47.826 (17.6)	22	5.580	.000
All Groups[a]	69.194 (58.538)	58.538 (23.123)	60	6.711	.000

Note: a) wwContext > SwContext

The results indicate that for all groups, the unstressed Function words occurring in ww contexts (i.e. following an unstressed syllable) were more accurately repeated than those occurring in Sw contexts (i.e. following a stressed syllable). This is the opposite result to that predicted by McGregor & Leonard's hypothesis. Further analysis of this finding suggests that Function word accuracy may be more influenced by broader sentence-level priming than the local prosodic context (see Section 8.14.3.2).

8.12 Effect of Sentence-Level Rhythm Regularity on Accuracy

We had hypothesised that a regular sentence rhythm might facilitate accurate repetition, resulting in higher scores for sentences with a regular rhythm than sentences with a variable rhythm. Half of the sentences therefore had a regular internal rhythm with a regular number of unstressed syllables (two) between stressed syllables. Half had a variable rhythm, where the number of unstressed syllables between stressed syllables could be between one and four.

Two new variables were created (Regular Rhythm and Variable Rhythm) by taking the mean of the percentage scores for each variable.

Table 8-27 Results of paired sample t-tests for each group comparing accuracy for regular and variable rhythm sentences

	Regular Rhythm Mean (SD)	**Variable Rhythm Mean (SD)**	***df***	***t***	***p***
Sentence Score Number correct					
AMC	1.479 (.824)	1.729 (1.05)	23	-2.061	.051
DLD	.107 (.162)	.071 (.153)	13	.618	.547
YLC	.130 (.249)	.098 (.146)	22	.768	.451
All Groups	.656 (.858)	.734 (1.044)	60	-1.432	.157
Word Score % Correct					
AMC[a]	80.017 (10.533)	85.963 (10.319)	23	-3.516	.002
DLD	45.573 (11.916)	48.853 (11.674)	13	-1.691	.115
YLC[a]	50.698 (16.545)	57.017 (14.576)	22	-3.079	.005
All Groups[a]	61.057 (20.36)	66.532 (20.172)	60	-4.947	.000
Syllable Score % Correct					
AMC	85.642 (11.599)	87.509 (9.985)	23	-.976	.339
DLD	49.263 (11.697)	49.874 (11.63)	13	-.310	.762
YLC	57.998 (17.702)	61.155 (15.117)	22	-1.491	.150
All Groups	66.87 (20.966)	68.934 (19.932)	60	-1.757	.084

Note: a) Variable > Regular

There was no significant effect of rhythm on the sentence or syllable scores for any of the groups. Conversely, there was a significant effect of rhythm on the word scores for the AMC and YLC groups, but no effect for the DLD group. For the AMC and YLC groups, the percentage of words repeated correctly was higher for the Variable rhythm sentences than for the Regular rhythm sentences. This was an unexpected result as it was in the opposite direction from our prediction.

To explore this word level result further, the words were divided according to grammatical category, and a repeated-measures ANOVA (3 x 3 x 2 – Group [AMC, DLD, YLC] x Grammatical Status [Content, Detaux, PrepConj] x Rhythm [Regular, Variable]) was conducted. There was a significant main effect of Rhythm *F*(1, 58) = 18.934, *p* = .000, and a significant Rhythm*Grammatical Status interaction *F*(1.594, 116) = 6.185, *p* = .003 (Greenhouse-Geisser correction).

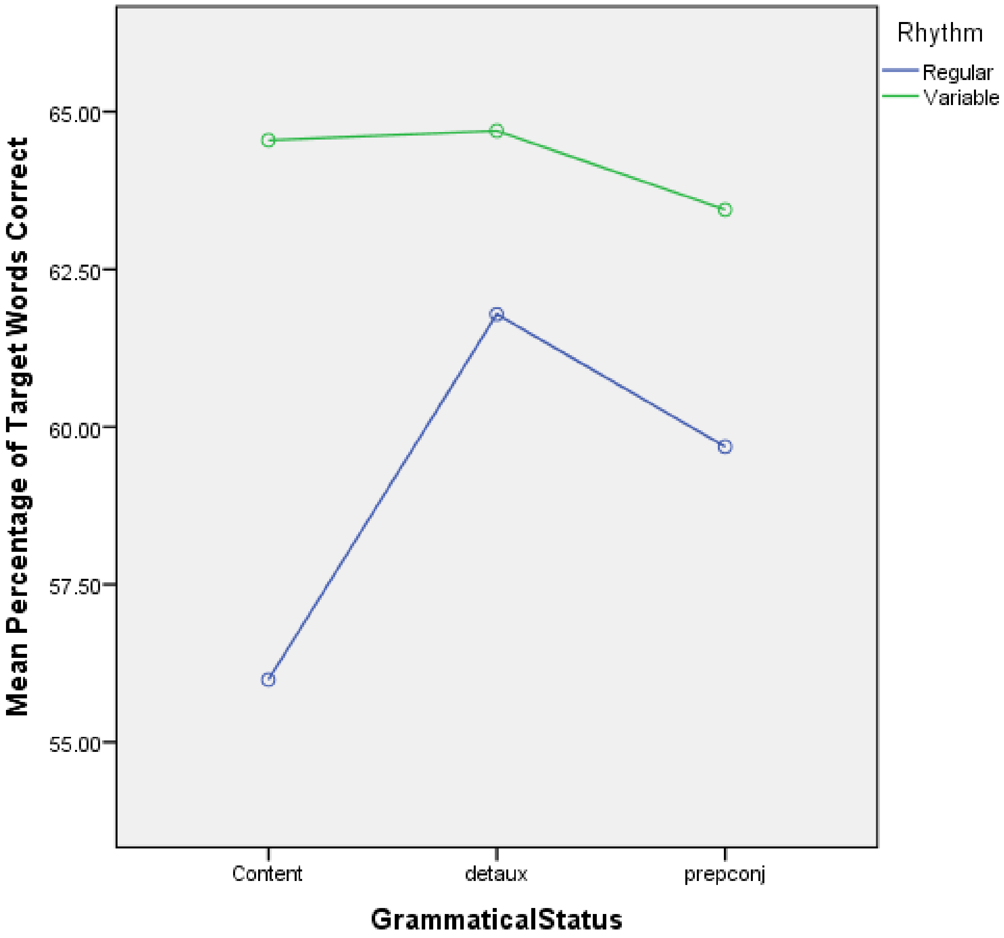


Figure 8-12 Graph showing mean percentage of target words correct by grammatical status for Regular and Variable rhythm sentences

From inspecting the graph (Figure 8-12), it seems that this interaction was driven by relatively poor Content word accuracy in Regular sentences compared with Variable sentences.

8.12.1 Summary of Effect of Sentence-Level Rhythmic Regularity

Regularity of the sentence-level rhythmic structure did not significantly affect repetition accuracy at either sentence or syllable level. A significant result was obtained at word level for AMC and YLC groups, with words within variable rhythms more accurate than those within regular rhythms. When examined by grammatical category, this effect was particularly driven by differences in Content word repetition and therefore seems to be unlikely to be an effect of the global rhythmic structure, but more likely to be caused by the local, internal prosodic structure of individual target words. Overall regularity of rhythm therefore does not seem to be a significant factor in accurate repetition.

8.13 Relationship of Sentence Repetition Scores to Acoustic Thresholds

We predicted that success in Sentence Repetition may be related to the children's scores in the Acoustic Threshold Estimation tasks (Chapter 4). We thought that children who had greater sensitivity to small differences in the acoustic measures (i.e. had lower thresholds) may be more accurate in their repetitions. In particular, based on our previous work, we predicted that repetition scores would be related to the Rise Time, Frequency and Duration measures.

Correlation co-efficients were therefore calculated using the accuracy scores at sentence, word and syllable level.

Table 8-28 Results of Correlations (Pearson one-tailed) AMC and DLD groups between Acoustic Threshold and Sentence Repetition Accuracy measures

	Duration	Frequency	Intensity	Sentence Score	Word Score	Syllable Score
Rise Time	.340*	.795***	-.227	-.510**	-.478**	-.476**
Duration		.406*	.016	-.416**	-.382**	-.394**
Frequency			-.206	-.614***	-.656***	-.643***
Intensity				-.010	.052	.087
Sentence Score					.855***	.830***
Word Score						.992***

Note: **p* < .05; ***p* < .01; ****p* < .001

Frequency thresholds were highly correlated with scores at sentence, word and syllable- level (p = .000) as were Rise Time thresholds ($p < .01$). Duration thresholds also correlated with Repetition performance, at sentence level ($p <.01$) and word and syllable levels ($p < .05$). For all three acoustic measures, the negative co-efficient indicates that the smaller the threshold (i.e. greater sensitivity to small changes), the better the repetition score. Intensity thresholds did not correlate significantly with repetition accuracy scores.

A set of regression equations was conducted to explore the unique variance in Sentence Repetition scores accounted for by the predictors of Age, NVIQ and each AT measure in turn. Step 1 was always entered as Age (months), Step 2 as NVIQ and Step 3 as the AT measure. Separate analyses were conducted with Sentence Score, Word Score and Syllable Score as the dependent variables.

Table 8-29 Results of Regressions exploring the unique variance in Sentence Score accounted for by Age (months), NVIQ and AT measures (AMC and DLD groups)

Sentence Score	**b**	**SEb**	**β**	***t***	***p***	**ΔR^2**	***p***
Model 1							
Age	.360	.182	.238	1.984	.056	.191	.007
NVIQ	6.227	1.398	.533	4.455	.000	.343	.000
Rise Time	-.075	.037	-.251	-2.021	.051	.051	.051
Model 2							
Age	.401	.180	.265	2.232	.033	.191	.007
NVIQ	6.371	1.405	.545	4.535	.000	.343	.000
Duration	-.108	.060	-.218	-1.794	.082	.041	.082
Model 3							
Age	.266	.199	.169	1.334	.192	.195	.008
NVIQ	6.095	1.633	.503	3.732	.001	.345	.000
Frequency	-11.490	5.814	-.275	-1.976	.057	.052	.057
Model 4							
Age	.470	.180	.323	2.616	.013	.186	.009
NVIQ	6.817	1.360	.607	5.014	.000	.359	.000
Intensity	-1.564	1.961	-.096	-.797	.431	.009	.431

Table 8-30 Results of Regressions exploring the unique variance in Word Score accounted for by Age (months), NVIQ and AT measures (AMC and DLD groups)

Word Score	**b**	**SEb**	**β**	***t***	***p***	**ΔR^2**	***p***
Model 1							
Age	.275	.149	.242	1.847	.074	.181	.009
NVIQ	4.111	1.145	.469	3.589	.001	.273	.000
Rise Time	-.055	.030	-.246	-1.815	.079	.050	.079
Model 2							
Age	.297	.145	.261	2.049	.048	.181	.009
NVIQ	4.137	1.132	.472	3.655	.001	.273	.000
Duration	-.093	.049	-.250	-1.916	.064	.055	.064
Model 3							
Age	.190	.158	.161	1.206	.237	.188	.009
NVIQ	3.510	1.294	.386	2.713	.011	.264	.000
Frequency	-11.596	4.607	-.370	-2.517	.017	.093	.017
Model 4							
Age	.356	.151	.321	2.357	.025	.175	.011
NVIQ	4.580	1.143	.534	4.008	.000	.278	.000
Intensity	-.916	1.648	-.074	-.556	.582	.005	.582

Table 8-31 Results of Regressions exploring the unique variance in Syllable Score accounted for by Age (months), NVIQ and AT measures (AMC and DLD groups)

Syllable Score	**b**	**SEb**	**β**	***t***	***p***	**ΔR²**	***p***
Model 1							
Age	.280	.151	.247	1.858	.072	.183	.008
NVIQ	3.957	1.161	.452	3.410	.002	.256	.000
Rise Time	-.055	.031	-.247	-1.796	.082	.050	.082
Model 2							
Age	.298	.146	.263	2.049	.049	.183	.008
NVIQ	3.945	1.138	.451	3.467	.001	.256	.000
Duration	-.100	.049	-.268	-2.043	.049	.063	.049
Model 3							
Age	.210	.161	.178	1.298	.204	.193	.008
NVIQ	3.316	1.324	.366	2.504	.018	.240	.001
Frequency	-11.205	4.715	-.359	-2.376	.024	.087	.024
Model 4							
Age	.354	.154	.318	2.295	.028	.177	.011
NVIQ	4.454	1.166	.519	3.820	.001	.260	.000
Intensity	-.434	1.682	-.035	-.258	.798	.001	.798

Since Frequency and Duration both added significant unique variance (8.7% and 6.3%) with further additional variance contributed by Rise Time (5%) when added at Step 3, a further model was created to further examine the influence of each AT variable. Frequency (with the highest amount of unique variance) was added at Step 3, followed by Duration and then Rise Time.

Table 8-32 Results of Regressions exploring the unique variance in Syllable Score accounted for by Age (months), NVIQ and AT measures (AMC and DLD groups)

Syllable Score	**b**	**SEb**	**β**	***t***	***p***	**ΔR^2**	***p***
Model 5							
Age	.203	.163	.172	1.240	.225	.193	.008
NVIQ	3.140	1.346	.346	2.332	.027	.240	.001
Frequency	-9.596	7.474	-.307	-1.284	.209	.087	.024
Duration	-.063	.054	-.165	-1.161	.255	.022	.243
Rise Time	.003	.048	.012	.059	.954	.000	.954

The pattern of results was similar across Sentence, Word and Syllable analyses. Both Age and NVIQ accounted for significant amounts of unique variance (Age variance range: 17.5% - 19.5%, *p* range = .007 - .011; NVIQ variance range: 24% - 35.9%, *p* range = .000 - .001). The contribution of Rise Time (range 5% - 5.1%) did not reach significance (*p* range = .051 - .082), whilst Duration was a significant predictor at Syllable level only (6.3% variance, *p* = .049). Frequency contributed further significant variance at Word (9.3%, *p* = .017) and Syllable (8.7%, p = .024) levels. Intensity did not contribute significantly to the accuracy scores at any level. The additional regression at Syllable level (Model 5) revealed that of the co-varying factors of Frequency and Duration, the contribution of Duration (2.2%) at Step 4 was no longer significant (*p* = .243) when Frequency was entered into the model at Step 3.

We had predicted that Frequency, Rise Time and Duration measures would be related to Sentence Repetition performance, and we found a significant correlation for scores at sentence, word and syllable level. For all three measures, greater sensitivity to acoustic change was related to higher repetition accuracy. Intensity thresholds did not, however, relate to repetition performance. When examining the individual contributions made by the three co-varying cues, Frequency was revealed to contribute the most unique variance above Age and NVIQ.

8.14 Results: Thematic Analysis

The statistical analysis of the overall dataset lent some support to the idea that categorical grammatical status may have some influence on repetition at certain grain-sizes for AMC and YLC groups as well as to the idea that prosody influences repetition accuracy with a privileged status for stressed syllables. It could therefore be that the best explanation of accuracy patterns lies not in looking individually for whether grammatical or prosodic features dominate responses, but in looking at the interaction of syntactic and prosodic levels. In this section, therefore, a more holistic analysis of the sentences will be presented, focusing on the similarities and differences in accuracy patterns across the eight different prosodic-syntactic sentence types used in the task.

8.14.1 Compiling the data

In order to explore the data from this different perspective, a score chart was constructed syllable-by-syllable for each target sentence (Total accuracy is expressed as the summed percentage correct from all three groups, AMC, DLD and YLC, i.e. maximum score of 300 per syllable). These were then summed across the four exemplars of each sentence type to create a maximum cumulative percentage score of 1200 per syllable. Since there was no general effect of entrainment in the statistical analysis, both entrained and unentrained scores were used for each sentence. Examining the patterning of the resultant charts allowed for the exploration of syntax and prosody in a multi-dimensional way by reflecting both the prosodic and syntactic status of syllables as distributed in the sentence rather than in the more decontextualised categorisation of the overall statistical analysis.

8.14.2 Exploring the Embedded Sentence Structure

Four of the eight sentence-types had the same embedded sentence syntactic structure. Within that framework, the prosodic structure of the Content words varied in order to create different prosodic patterns – see Table 8-33.

Table 8-33 Example target sentences for each of the four embedded sentence constructions

Row	Type	Example Sentence												Rhythm
1	**Emb-Reg-Sw**	the	**CAR**	pet	the	**ME**	tal	is	**UN**	der	is	**PUR**	ple	regular
2	**Emb-Reg-nonSw**	the	dis	**PLAY**	the	gui	**TAR**	is	be	**HIND**	is	com	**PLETE**	regular
3	**Emb-Var-Sw**	the	**TA**	bles	the	com	**PU**	ter	is	**UN**	der	are	**BIG**	variable
4	**Emb-Var-nonSw**	the	ma	**CHINE**	the	ma	ga	**ZINE**	is	be	**HIND**	is	**SHORT**	variable

Note: dark shading indicates S(w) words, light shading indicates nonSw words

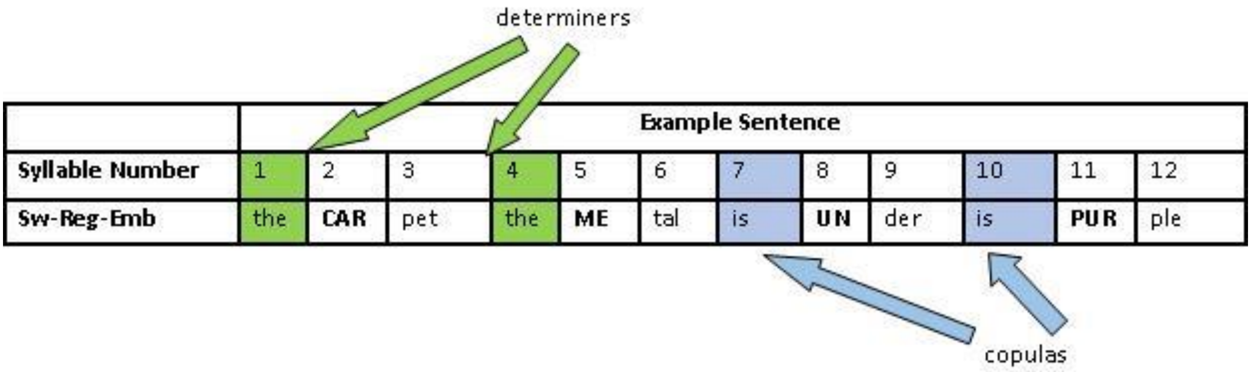
Two of the sentence types had a regular prosodic structure (Rows 1 and 2) and for the other two (Rows 3 and 4), the prosodic pattern was variable.

Furthermore, two sentence-types had exclusively or predominantly Sw Content words (Row 1: Emb-Reg-Sw; Row 3: Emb-Var-Sw) and two sentence-types with exclusively or predominantly wS, wSw or wwS Content words (Row 2: Emb-Reg-nonSw; Row 4: Emb-Var-nonSw). This patterning was not itself a deliberate experimental manipulation, but was a by-product of the considerations of rhythm regularity and prosodic context for unstressed function words. Nonetheless, as we shall see throughout this section, this variation between predominantly Sw and nonSw stress patterns proved to be of interest. There were four target sentences for each sentence type.

A syllable-by-syllable accuracy chart was created for each of the four constructions and then examined for grammatical and prosodic influences.

8.14.2.1 Considering the role of grammar

The original consideration in creating the sentence types was to compare accuracy across Function and Content words. The Function words in the embedded sentences were determiners in the noun phrases, and copulas introducing the prepositional phrase (embedded clause) and adjective phrase (main clause) – see Figure 8-13.



	Example Sentence											
Syllable Number	1	2	3	4	5	6	7	8	9	10	11	12
Sw-Reg-Emb	the	CAR	pet	the	ME	tal	is	UN	der	is	PUR	ple

Figure 8-13 Example sentence illustrating location of determiners and copulas in Embedded Sentence structures

The accuracy rates for the determiners and copulas are highlighted in Figures 8-14 and 8-15 for the Regular and Variable embedded sentences.

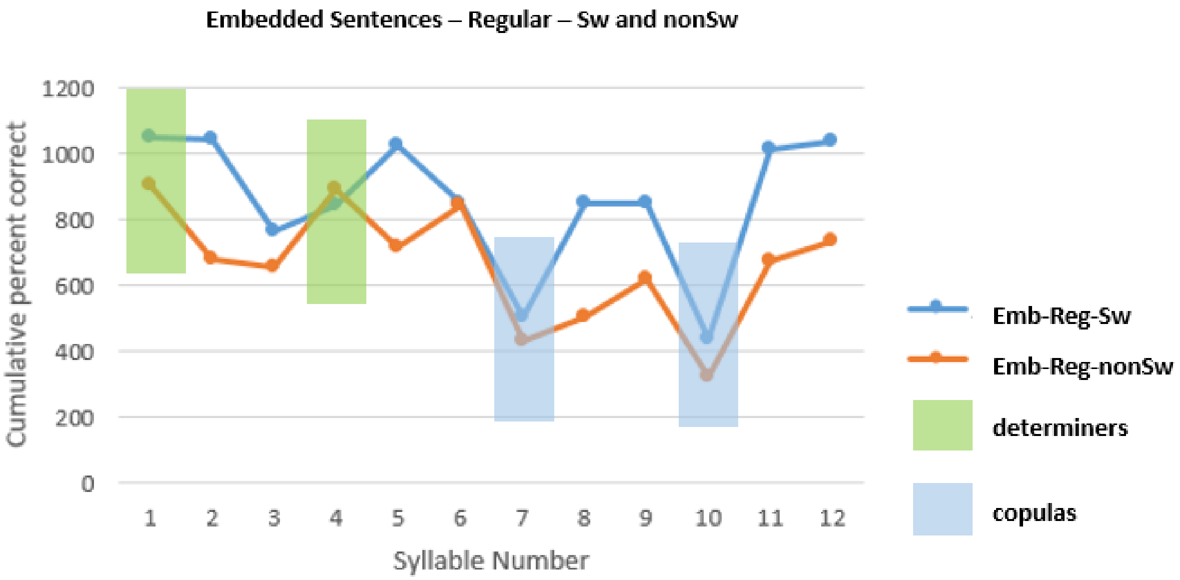


Figure 8-14 Accuracy graph for and Emb-Reg-Sw and -nonSw sentences with location of determiners and copulas highlighted

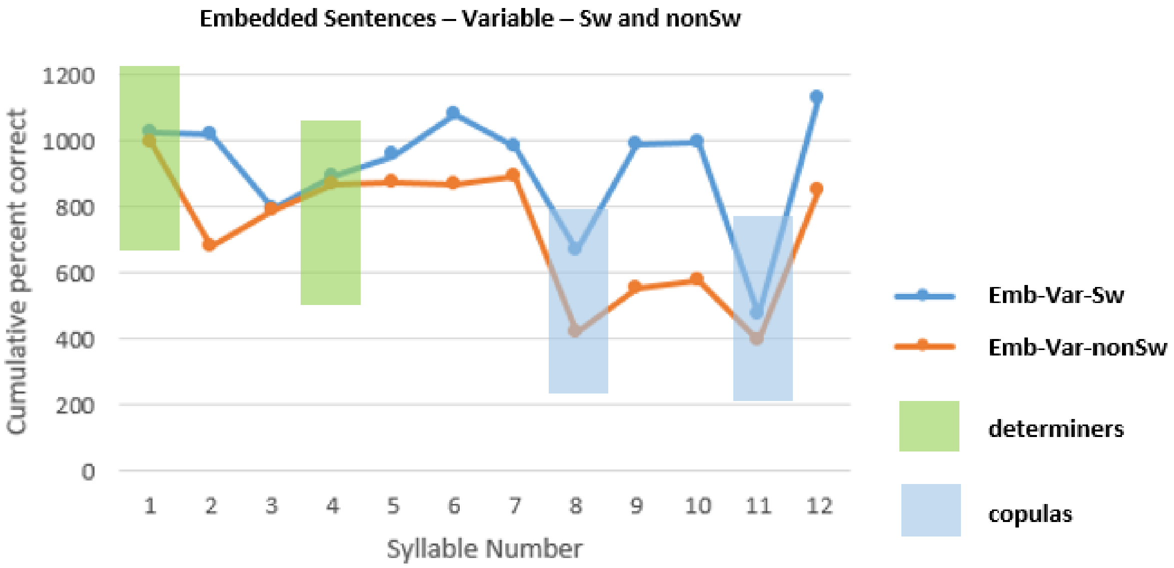


Figure 8-15 Accuracy graph for Sw-and Emb-Var-nonSw sentences with location of determiners and copulas highlighted

By visually comparing the accuracy rates for the determiners (green shading) and copulas (blue shading), we can see that the notion of 'Function' is insufficient to explain the response pattern. In fact, the determiners appear to show a relatively good level of accuracy, whilst there appears to be a generally lower rate of accuracy for copulas. Having already noted that 'Function' masks some some important differences between Detaux and PrepConj words and syllables in the statistical analysis, it seems as though a further division of the Detaux category into determiner and copula syllables may also be necessary in order to more accurately capture accuracy patterns.

At a phrasal level, the situation also differs between the noun phrases and copula structures. Within the noun phrases there seems little overall differentiation in accuracy between the determiner and its noun, whereas for the copula structures, the preposition/adjective element appears markedly more accurate than the copula 'is'.

At a sentence level, the effect of syntactic structure on accuracy can be most clearly seen if we chart the two Sw sentence-types together (thus minimising any prosodic variation).

Table 8-34 Example Sw sentences indicating the location of noun and copula structures

Type	Syllable Number											
	1	2	3	4	5	6	7	8	9	10	11	12
Emb-Reg-Sw	The	**CAR**	pet	the	**ME**	tal	is	**UN**	der	is	**PUR**	ple
Emb-Var-Sw	The	**TA**	bles	the	com	**PU**	ter	is	**UN**	der	are	**BIG**

Note: blue shading indicates noun phrases; pink shading indicates copula structures

We can see from Figure 8-16 that the accuracy scores for the opening noun phrase (syllables 1-3) coalesce in accuracy, diverging over syllables 4-6 in consequence of the differing prosodic structures of the second noun phrase. Following that, there is a patterned divergence as the syntactic structure begins the copula structure at syllable 7 in Emb-Reg-Sw (blue line), then a parallel effect when the copula structure begins at syllable 8 in Emb-Var-Sw (purple line). The remainder of the lines then continue to mirror each other at one syllable's distance.

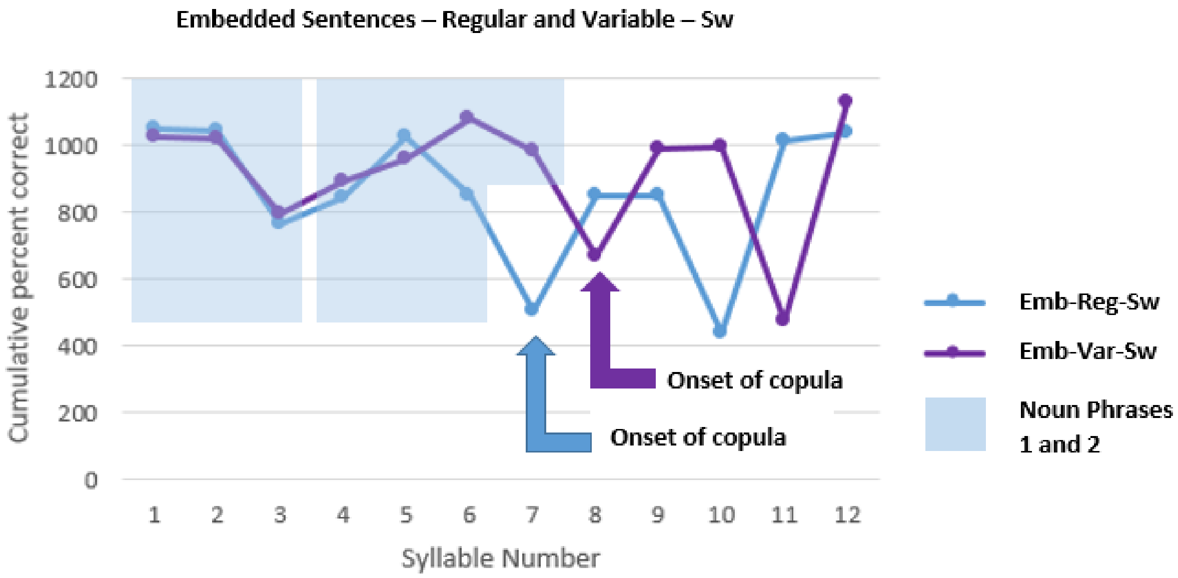


Figure 8-16 Accuracy graph for Embedded sentences-Regular and Variable-Sw highlighting location of copula structures

The parallel accuracy rates coinciding with the parallel syntactic structures of the two sentence-types suggest that sentence structure is directly influencing the accuracy of repetition.

Grammar cannot tell the whole story, however. If response accuracy were predicted on the basis of grammar alone, then we would expect the line of all four sentence-types to mimic the same path regardless of the prosodic structure, since each sentence has identical syntax. The deviations in accuracy between the Sw and nonSw sentence types as seen in Figures 8-14 and 8-15 illustrate that this is not the case, as the line for each sentence type follows a divergent course within the overall grammatical structure. We will therefore next consider what influence prosody may have on the results.

8.14.2.2 Considering prosodic variations

The original intention behind the creation of the prosodic structure of the sentences was to produce two different prosodic contexts for Detaux category words. In the Sw sentences (Rows 1 and 2 of Table 8-35, below), Detaux words occur following a weak or unstressed syllable (highlighted in purple). In the nonSw sentences, Detaux words occur following a strong syllable (highlighted in yellow).

Table 8-35 Examples of Embedded sentence stimuli with Detaux words highlighted

Row	**Type**	**Example Sentence**												**Rhythm**
1	**Emb-Reg-Sw**	the	**CAR**	pet	the	**ME**	tal	is	**UN**	der	is	**PUR**	ple	regular
2	**Emb-Var-Sw**	the	**TA**	bles	the	com	**PU**	ter	is	**UN**	der	are	**BIG**	variable
3	**Emb-Reg-nonSw**	the	dis	**PLAY**	the	gui	**TAR**	is	be	**HIND**	is	com	**PLETE**	regular
4	**Emb-Var-nonSw**	the	ma	**CHINE**	the	ma	ga	**ZINE**	is	be	**HIND**	is	**SHORT**	variable

Note: purple shading = Detaux word following unstressed syllable; yellow shading = Detaux word following stressed syllable

If we revisit the graphs of Figures 8-14 and 8-15, this time with the target Detaux words highlighted (Figures 8-17 and 8-18 below), we can see that there are minor differences in accuracy for individual Detaux words in the majority of cases, whether the preceding noun is Strong (as for the yellow line) or weak (purple line).

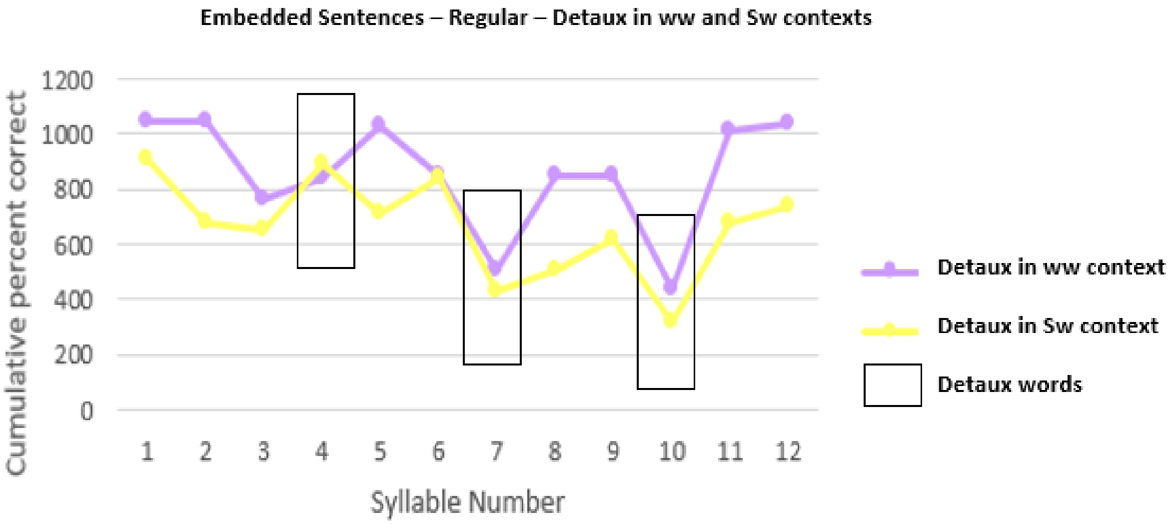


Figure 8-17 Accuracy graph for regular rhythm embedded sentences with location of Detaux words highlighted

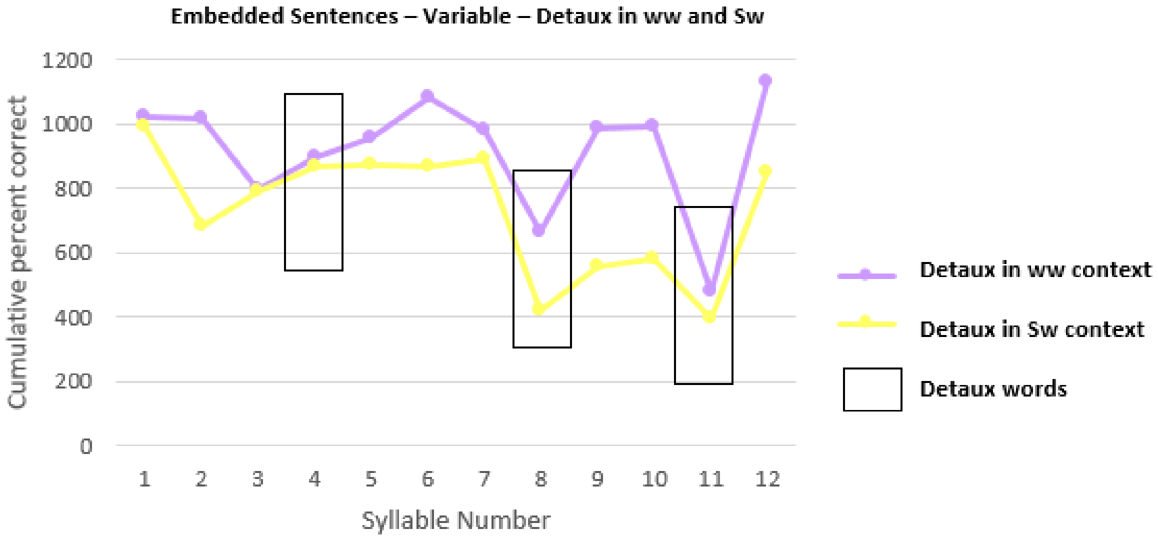


Figure 8-18 Accuracy graph for variable rhythm embedded sentences with location of Detaux words highlighted.

The prosodic theory put forward by McGregor & Leonard (1994) would have predicted that when an unstressed Function word occurs directly after a preceding stressed syllable it is more likely to be preserved, however the trend observable is for the wwContext Function words to be slightly more accurate.

Having explored the originally predicted prosodic effect, we looked for any prosodic influences at sentence and lexical level. Revisiting Figures 8-14 and 8-15 to consider accuracy levels across entire sentences, we can see that the Emb-Reg-Sw sentence-type is generally more accurate than its counterpart Emb-Reg-nonSw sentence-type, and that similarly, the Emb-Var-Sw sentence-type is generally more accurate than the Emb-Var-nonSw sentence-type.

These are sentence pairs with identical distribution of phrase and syntax across the syllables, and thus this difference cannot be explained by syntax alone. The most likely explanation is therefore in their differing prosodic structures, with those sentences containing predominantly Sw patterns within the lexical prosodic structure (e.g. CARpet) being more accurate than those whose lexical prosodic structure is predominantly nonSw (e.g. guiTAR; comPUter). This suggests that sentences containing predominantly nonSw words are less accurately repeated overall than those containing predominantly Sw words, despite equivalent phrasal and sentence level syntax.

Prosodic influence is also evident at lexical and syllabic levels. Figures 8-14 and 8-15 are reproduced again below (as Figures 8-19 and 8-20), but this time with the different nouns, prepositions and adjectives highlighted.

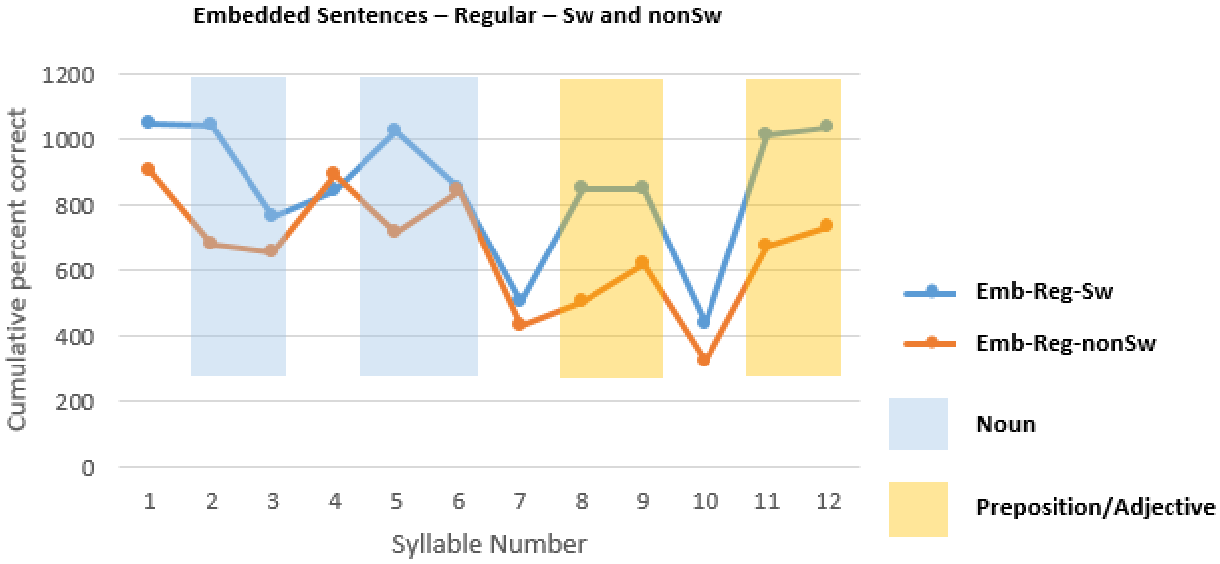


Figure 8-19 Accuracy graph for regular rhythm embedded sentences with location of nouns, prepositions and adjectives highlighted.

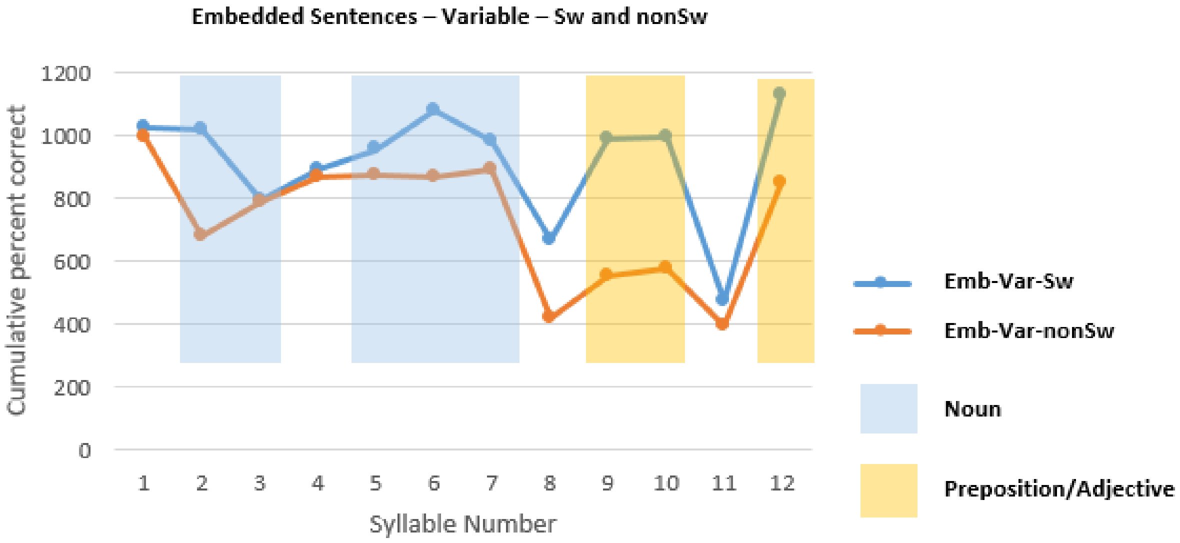


Figure 8-20 Accuracy graph for variable rhythm embedded sentences with location of nouns, prepositions and adjectives highlighted

Looking at the patterns for each highlighted word, we can see that syllable-level accuracy across the word is determined by its prosodic structure. To illustrate this, Figure 8-21 (overleaf) highlights syllables 2 & 3 and 5 & 6 from the Sw- and Emb-Reg-nonSw sentences. In each sentence-type these syllables represent the two nouns.

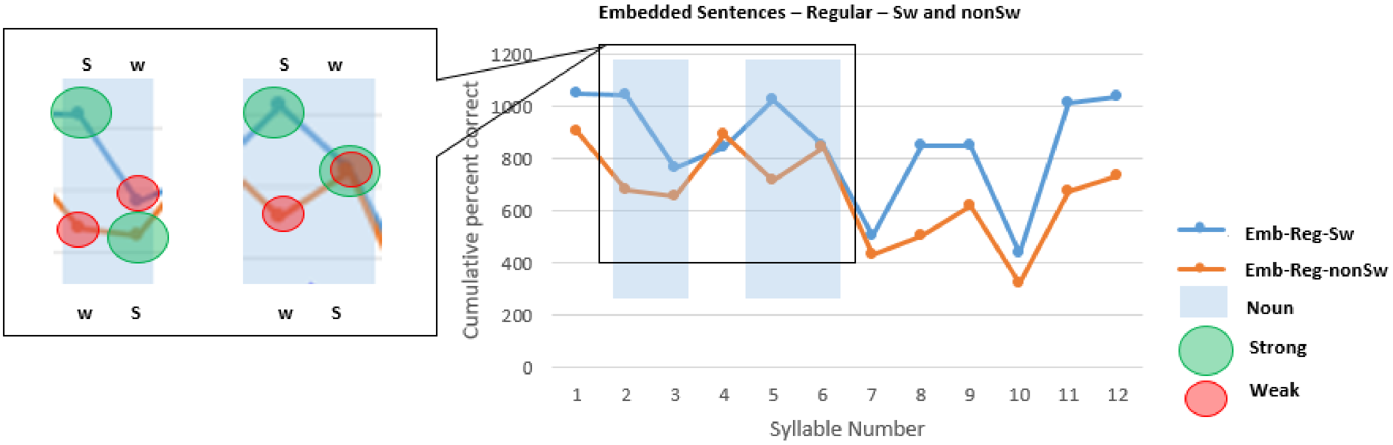


Figure 8-21 Accuracy graph of the Regular rhythm embedded sentences with syllables 2 & 3 and 5 & 6 highlighted

In the Emb-Reg-Sw sentences, the two target nouns have a Sw pattern (e.g. MEtal, CARpet) whereas for then Emb-Reg-nonSw sentences, these are both wS nouns (e.g. disPLAY, reWARD). The tendency for the stressed syllable of each word to have a greater relative accuracy is illustrated in the call-out box of Figure 8-21. Furthermore, despite the generally lower accuracy of the Emb-Reg-nonSw sentence-type, there is still a relative advantage for the stressed syllables within each target word across the sentence, seen particularly in syllables 6, 9 and 12 in the above example (Figure 8-21).

If the prosodic influence we can see here were generally the case, we would also expect to see a similar pattern emerging in the accuracy patterns of the Emb-Var-Sw and Emb-Var-nonSw sentence-types, which again vary only in prosody and not in syntactic structure. Figure 8-22 highlights the first two noun phrases for these variable rhythm sentences.

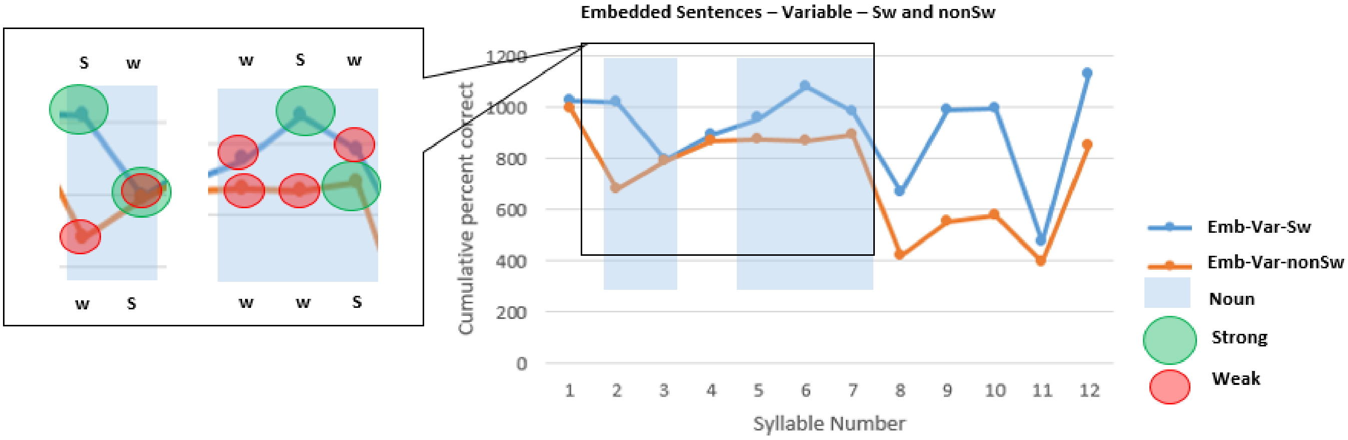


Figure 8-22 Accuracy graph of variable rhythm embedded sentences with syllables 2 & 3 and syllables 5, 6 &7 highlighted

Inspecting the graphs of these two noun phrases, there is again a discernible reflection of the prosodic structure in the accuracy rates across syllables. The pattern for the first noun mirrors that we saw in the previous illustration with the difference between the Sw and wS nouns clearly visible. The second noun also reflects the different prosodic structures of the wSw target in the Emb-Var-Sw (blue) line and the wwS target of the Emb-Var-nonSw (orange) line. Again, the stressed syllable is reflected in a relative increase in accuracy within the target word.

A notable feature across these figures is that the stressed syllable accuracy advantage is relative, i.e. within the context of an individual word, not an absolute value across the sentence. To take an example, Figure 8-23 highlights two words from the Emb-Reg-nonSw sentence-type (orange line). Syllables 5 & 6 form a noun (e.g. guiTAR) and syllables 8 & 9 form a preposition (e.g. beHIND).

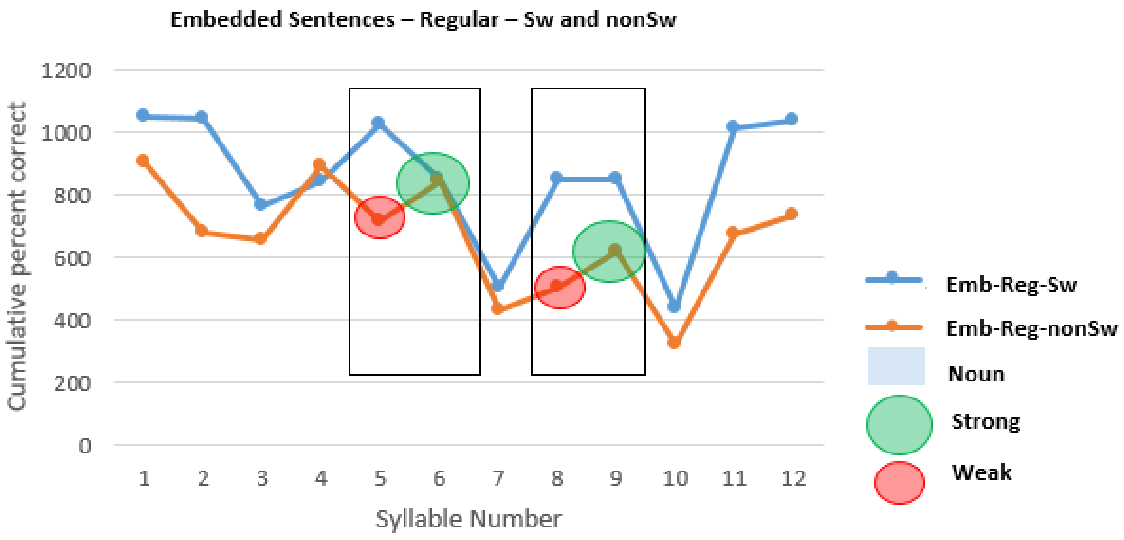


Figure 8-23 Accuracy graph of Regular rhythm embedded sentences with Syllables 5 & 6 and 8 & 9 highlighted

The weak syllables (e.g. 5 "gui" and 8 "be") are clearly less accurate than their strong counterparts (e.g. 6 "TAR" and 9 "HIND"). This is what we would have expected from the strong statistical result privileging stressed over unstressed syllables. However, by looking at the words within the context of the whole sentence, we can also see that the strong syllable 9 is substantially less accurate than strong syllable 6, and, importantly, also less accurate than weak syllable 5. We can also see that the strong syllable 6 of the wS words in the Emb-Reg-nonSw (orange line), is only as accurate as the weak syllable 6 of the equivalent Sw words in the Emb-Reg-Sw (blue line).

Stress and prosody therefore do not appear to be absolute constructs, but factors that operate on a relative scale, contributing relative accuracy values for a given syllable. This relativity appears not to operate at the level of the whole sentence, but at a lexical or phrasal level, with the stressed syllable advantage for one phrasal unit not necessarily directly equivalent to that of a second unit, even from within similar rhythmic and syntactic structures. The difference in absolute accuracy between targets of equivalent prosodic structure may be a reflection of the grammatical role that target phrase is playing in the sentence, and thus a reflection of the interaction of prosodic and syntactic levels.

Whilst the initial predictions regarding post-stress Function words were not supported by the data, there was nonetheless a prosodic influence on accuracy at sentence, phrasal and syllabic levels. Sentences rich in Sw structures are more accurate overall than those in which nonSw structures predominate. Within phrases, the internal prosodic structure is reflected in accuracy rates with a privileging of the stressed syllable. The prosodic advantage is a relative one, operating at a phrasal level which is internally consistent, but may differ in absolute value across phrases.

Combining these observations on the role of prosody with our findings of clausal and phrasal level grammatical influence, it seems that accuracy patterns across sentences are unlikely to be explained purely in terms of Function versus Content words, or stressed versus unstressed syllables, but instead reflect a complex interplay between different levels of linguistic structure at the prosodic, lexical, phrasal and clausal level.

8.14.3 Exploring the Complex Sentences

For the preceding interpretation of the accuracy patterns to prove robust, it should be reflected in other syntactic structures. We therefore also looked at the syllable-by-syllable accuracy of the 15-syllable complex sentences that formed the other half of the stimuli set. These were again all composed of identical syntactic structures, in this case consisting of two clauses, each composed of a noun phrase and a verb phrase, centrally linked with a subordinating conjunction.

Table 8-36 Example sentences of the four complex sentence constructions

Type	**Example Sentences**														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Com-Reg-Sw	the	**STU**	dents	were	**AR**	gu	ing	**AF**	ter	the	**MEE**	ting	had	**STAR**	ted
Com-Reg-nonSw	the	ho	**TEL**	was	co	**LLAP**	sing	be	**FORE**	the	po	**LICE**	had	pre	**PARED**
Com-Var-Sw	the	ex	**PLO**	rer	was	**RES**	ting	**AF**	ter	the	ad	**VEN**	ture	had	**STOPPED**
Com-Var-nonSw	the	**TEAM**	were	**CE**	le	bra	ting	be	**FORE**	the	**MATCH**	had	con	**CLU**	ded

As with the embedded sentences, two of the sentence-types had a regular and consistent prosodic structure (one predominantly Sw – Row 1, one predominantly wS – Row 2). The second two sentence-types had an irregular prosodic structure which varied between exemplars (Rows 3 and 4).

8.14.3.1 Considering grammatical variation in Complex Sentences

We began by considering whether the grammatical influences that we saw with Embedded Sentences were also present in the accuracy graphs for the complex sentences.

Because the prosodic structure of the Variable Rhythm sentence-types varied across their component sentences, and so syllable 13, for example, may be an unstressed auxiliary syllable in one sentence and a stressed Content syllable in another, it is more useful to illustrate the more detailed patterning with the Regular Rhythm sentences only, where the prosodic and grammatical patterning is more directly comparable in graphical form.

Firstly, looking at the overall sentence structure in Figure 8-24, there is a pattern to the accuracy levels, pivoting around the conjunction (syllables 8-9). There is a fairly steady level of accuracy across the main clause (with specific variation as we will discuss below) then a sharp tail-off in accuracy over the course of the subordinate clause. The specific clausal structures therefore appear to be differentially influencing accuracy levels.

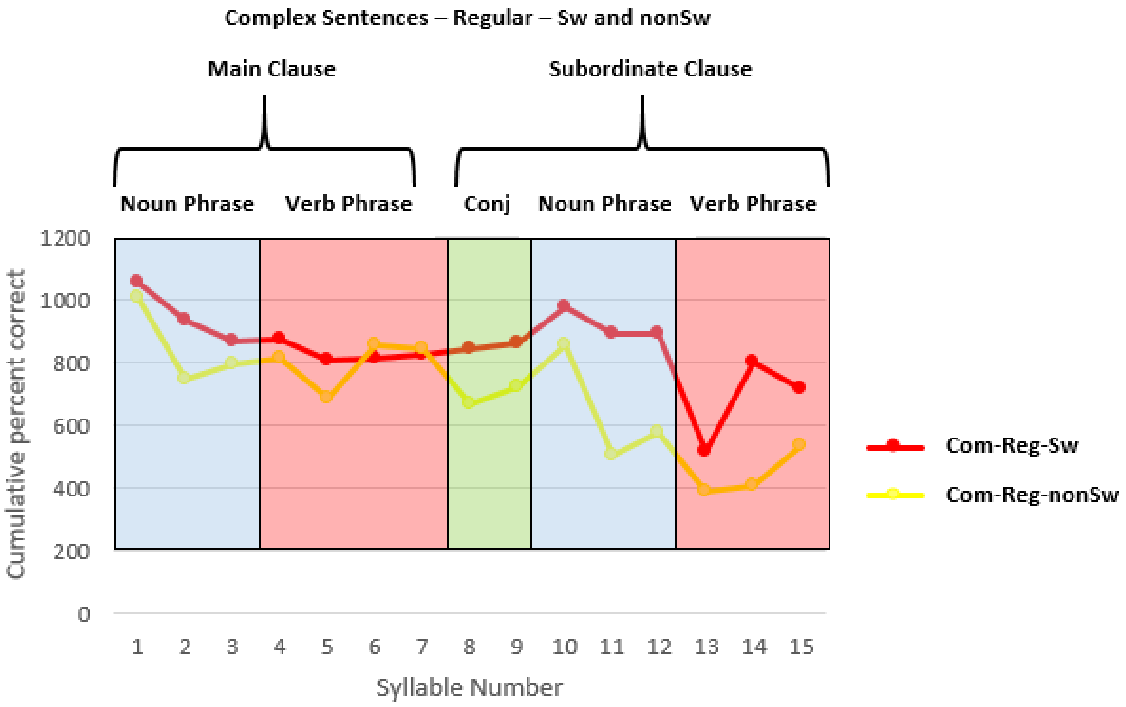


Figure 8-24 Accuracy graph for Regular Complex sentences with the component syntactic structures highlighted

In the embedded sentences, we saw a differentiation between the accuracy of determiners and those of the copulas. In the subordinating sentences, the equivalent unstressed Function words were determiners in the noun phrases and auxiliaries in the verb phrases.

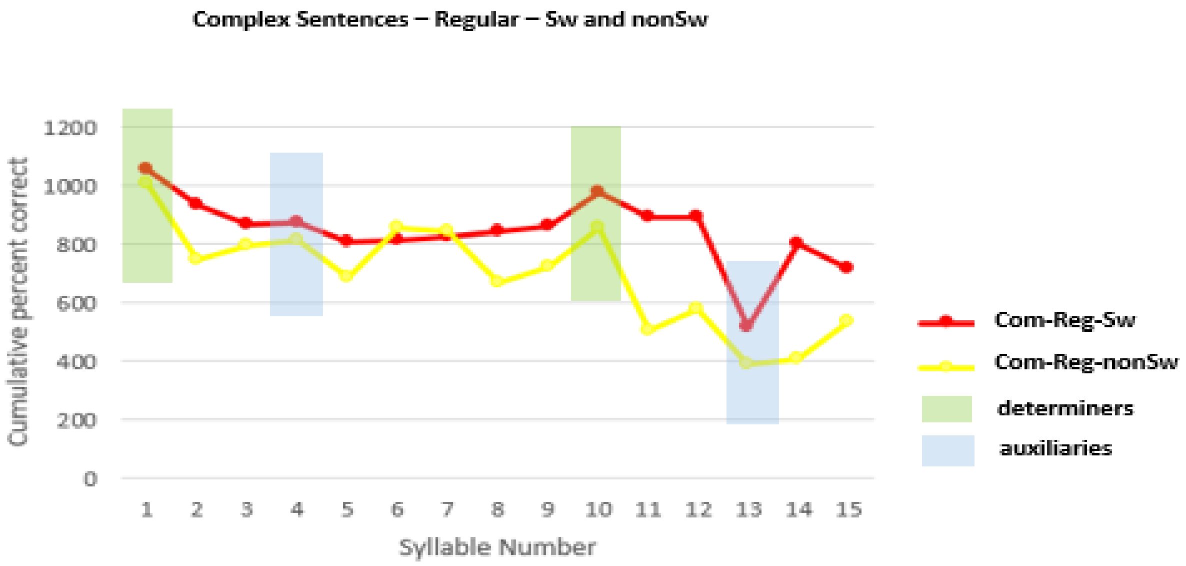


Figure 8-25 Accuracy graph of Regular Complex Sentences with location of determiners and auxiliaries highlighted

At the phrasal level, we can see from Figure 8-25 that the determiners of the noun phrases (syllables 1 and 10) are relatively accurate in both main and subordinate clauses for both sentence types. There is a clear dip in accuracy for the final auxiliary (syllable 13), particularly noticeable in the Com-Reg-Sw sentence-type. This general trend cannot be explained by prosody and so is most likely a reflection of the structural properties of the sentence. In these sentences, there are two target auxiliaries: syllables 4 and 13. There is a distinction in accuracy between the two, with syllable 4 tending to have higher accuracy levels than 13. This suggests again that ‘Function’ is insufficient as an explanation, or even a sub-category of auxiliary

The patterning evident in these observations indicates that the grammatical structure of a sentence continues to influence accuracy rates of repetition for the complex sentences as well as the embedded sentences. There appears to be a relatively consistent relationship between determiner and noun within noun phrases and the accuracy of auxiliaries in verbal constructions may be influenced by familiarity of specific tense and construction requirements.

8.14.3.2 Exploring Prosodic Influences on Complex Sentences

We began by examining the predicted influence of post-stress positioning on Function words. Again, there appeared to be little difference between the accuracy of Function words which followed a stressed syllable compared with those which followed an unstressed syllable. If anything, there appears to be a small advantage for the Function words occurring following a weak syllable (represented by the Com-Reg-Sw – red line in Figure 8-25). This is the opposite effect to that predicted by McGregor & Leonard, and may be explained by a sentence-level advantage for those examples which contain predominantly Sw words.

We saw with the Embedded sentences that the sentences which contained predominantly Sw words were more accurate overall than the nonSw sentences. If we revisit the complex sentences – see Figure 8-25 we can observe a similar phenomenon, with the accuracy line for the Com-Reg-Sw sentence-type (red) being generally higher than that of the Com-Reg-nonSw line (yellow). Because the syntax of the two sentence-types is equivalent, this cannot be caused by grammatical influences and it is therefore most likely that this is a result of the prosodic differences created by the differing sentence rhythms. This observation may explain an unexpected result in the Statistical Analysis (Section 8.7) which demonstrated that unstressed Function words were more accurate when they occurred following an unstressed Content syllable (wwContext) than a stressed syllable (SwContest). Those Function words which occurred in an immediate wwContext were all part of Sw-Reg or Sw-Var sentences. It could therefore be this global rhythmic advantage that is supporting repetition of the weak Function words in these sentences, rather than the influence of local stress proximity as we had originally hypothesized.

At a lexical level, the Embedded Sentences demonstrated that the internal prosodic structure of the target words tended to be reflected in the relative accuracy of their component syllables.

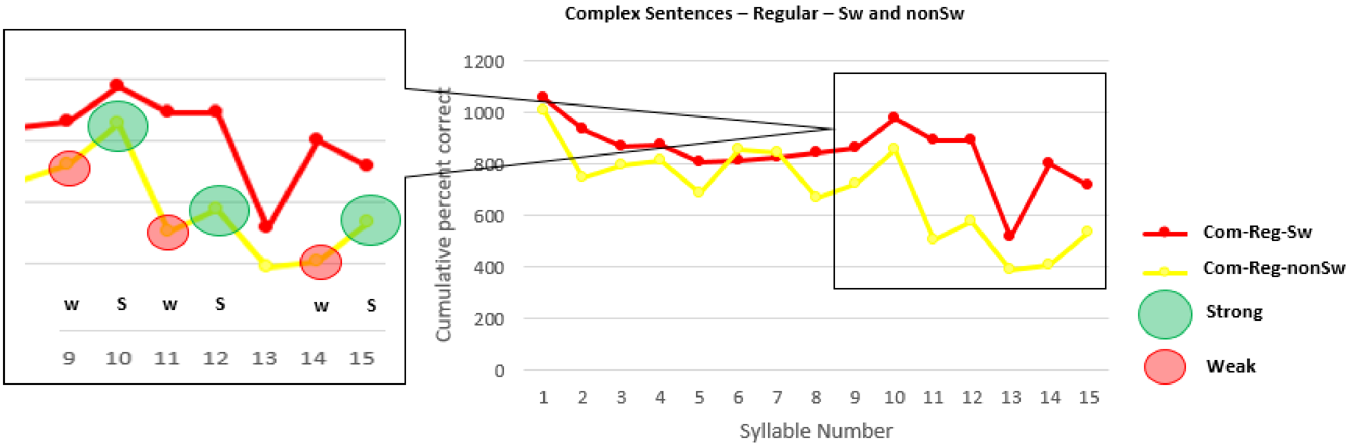


Figure 8-26 Accuracy Graph of Regular rhythm Complex sentences with the stress levels of words from the subordinate clause highlighted

Figure 8-26 illustrates a similar pattern occurring within the complex sentences as the highlighted weak-strong syllable patterns for the component words of the nonSw sentence-type (yellow line) demonstrate. Furthermore, the previously observed relative nature of this pattern remains evident – for example, the strong syllable 12 is substantially less accurate than weak syllable 9, whilst nonetheless preserving its relative accuracy level within its advantage over weak syllable 11 within its own lexical item.

We therefore find evidence for similar prosodic influences occurring in the complex sentences as we found with the embedded sentences, with prosodic structure reflected in accuracy levels at sentence, word and syllable levels.

8.14.4 Summarising the Findings of the Thematic Analysis

As we examined the potential influences of grammatical and prosodic structures across the two target syntactic structures, we found that similar themes emerged. There was a noticeable influence of syntax upon repetition accuracy. In the embedded sentences, the unfamiliarity of this construction caused accuracy problems for many children in creating the syntax of the final main clause. For the complex sentences, it seemed to be choice of auxiliary to create the correct tense in the subordinate clause which was most problematic. At a lexical level of grammar, we also see similar themes emerging to those we saw in the statistical analysis, with the division of words into ‘Function’ and ‘Content’ proving too simplistic to capture the overall patterns in the data. Whilst many of the Content words are repeated relatively accurately, so, frequently, are the prepositions, conjunctions and, indeed, the determiners. Conversely, the auxiliaries and copula verbs fare relatively poorly, but this in itself varies according to syntactic context with the imperfect auxiliaries more accurate on the whole than the pluperfect ones. It seems therefore that accuracy at the grammatical level is dependent on structural context rather than grammatical category per se.

We also see clear signs of prosodic influence on repetition accuracy. Sentences rich in Sw prosodic structures are repeated more accurately than those where nonSw structures predominate, despite equivalent syntax across sentences. This suggests a persistent processing advantage for the Sw patterns which are more prevalent in English as a whole. Within each phrase, the prosodic structure of the lexical item is reflected in the relative syllable-level accuracy, with a relative privileging of the stressed syllable. Conversely, we did not find evidence that those unstressed Function words which follow a stressed syllable are any better preserved than when they follow an unstressed one.

The observed accuracy patterns across the dataset therefore appear to be the result of an interaction of prosodic and syntactic structures. Within a given phrasal or lexical unit, the combined internal prosodic and syntactic structures appear to be producing relatively predictable accuracy patterns, but these are not reflected in absolute accuracy levels from one phrase to the next. Prosodic accuracy cueing therefore appears to be operating at neither sentence level nor on lexical units per se, but at the syntactic phrase level, where prosody and syntax combine to create relative patterns of error and accuracy.

8.15 Discussion

The task of sentence repetition, as has been recognised in previous literature (e.g. Polišenská et al., 2015), is a complex one, drawing upon a range of acoustic, linguistic and memory systems for successful completion. The sentence must first be processed acoustically, analysed linguistically with regard to phonology, syntax and semantics, then undergo redintegration of immediate memory traces with long-term knowledge to produce the response (e.g. Alloway & Gathercole, 2005).

The results presented here reflect this task complexity, with patterns consistent with all these layers of processing. Within this complexity, we were particularly focused on exploring the potential contribution of linguistic rhythm to task completion as this is an aspect of language structure that has not previously been investigated with regard to repetition tasks. We found that the rhythmic status or structure of the target syllable, word or sentence exerted an influence on how accurately that target was repeated; however this was not always in the predicted direction.

At a syllable level, there was a clear effect of stress status for the AMC and DLD groups with stressed syllables more accurately repeated than unstressed syllables. For the two grammatical categories containing both types of stress (Content and PrepConj), stressed syllables were more accurate than unstressed within each category. This result is consistent with the privileging of stress that we saw in the nonword repetition task, however here we can see that it also operates within words occurring as part of a larger linguistic structure.

At a word level, we were interested in further exploring the differing contributions of grammatical status (Content word or function word) and stress status (unstressed or carrying a stressed syllable). In previous work (e.g. McGregor & Leonard (1994)), all content words were monosyllabic and stressed; whilst all function words were monosyllabic and unstressed, leaving it unknown whether it was the grammatical category or the stress status which was the driver behind the result they obtained. Here, we deliberately chose function words which were unstressed (e.g. the, is) as well as those containing both a stressed and unstressed syllable (e.g. UNder, beHIND).

If the important distinction was the grammatical one between content and function, then we should find that accuracy rates for unstressed syllables differed between content and function words. We found that this was the case for the AMC group, but not for DLD or YLC groups, initially pointing to a role of grammatical category. Conversely, if the important distinction was the rhythmic one between words containing stress and words which did not contain stress, we should find that once the Function Word category was subdivided, there would be no difference between Content and PrepConj words (all of which contained a stressed syllable) but that unstressed syllables in these two categories differed from Unstressed syllables in Detaux words. This is what we found for the AMC group.

These scores suggest that it is the presence of stress which determines accurate repetition for the AMC group. When a stressed syllable is present in either Content or Function words, repetition of the corresponding unstressed syllables is more accurate than when an unstressed syllable stands alone (as in Detaux words). Further, there was no difference in accuracy between repetition of stressed syllables in Content and PrepConj words, supporting the equivalent role of stress across grammatical categories. These results support the role proposed for prosodic properties in accounts of the apparent vulnerability of function words (e.g. Gerken, 1996; McGregor & Leonard, 1994). Here, we deliberately chose function words containing primary stress, however function words with this property are in the minority, with some of the most frequent (a, the, is, was etc) being typically unstressed. The results obtained here for AMC children indicate that unstressed syllables occurring in isolation are more vulnerable than those occurring as part of a larger prosodic structure (e.g. a Sw or wS word).

Strikingly, there was no equivalent effect of stress status for the DLD group. One possible explanation may be in their relative insensitivity to the acoustic cues to rhythm and stress. Poorer sensitivity to stress cues may result in this group being unable to capitalise on the rhythmic properties of the target, in either creating or retrieving rhythmic templates in order to support their repetition. Since stress had an impact for this group at an individual syllable level (stress > unstressed) it could be that the lack of effect at word level may relate to a difficulty in integrating individual stress syllables into a larger prosodic structure i.e. in creating a representation of word-level patterns of stress. We must also account, however, for the lack of effect of stress status for the YLC group – these were younger, but typically-developing children. Perhaps this indicates that the development of robust rhythmic templates for words occurs across childhood, so that younger children with less linguistic exposure have yet to develop the full resources to draw on that the older typically-developing children were able to. This may leave them in a similar position to the older, DLD children.

Rhythmic structure was also seen to be operating at the level of the word. The graphs presented in the Thematic Analysis clearly show differing accuracy patterns for individual words depending on their internal prosodic structure. The results mimic those reported here and elsewhere in nonword repetition tasks (e.g Chiat & Roy, 2007) with weak syllables typically less accurate than strong syllables, and with wS structures typically less accurate overall than Sw structures. Such a result fits strongly with accounts such as Gerken's Metrical Template Hypothesis in which very young children have a bias towards producing structures which are consistent with a Sw template. Here, we see a similar pattern continuing to obtain with older, primary-school children.

Finally, the influence of rhythmic structure was seen at sentence level. This patterning was not deliberately experimentally manipulated, nevertheless it became clear that sentences rich in Sw structures were significantly more accurately repeated overall than those containing predominantly nonSw structures, despite syntax and regularity of sentence rhythm being held constant. For Content and PrepConj words, the differing internal prosodic structures could have accounted for this difference, however a significant difference was also found for Detaux words. Detaux words were unstressed in all sentence types, and since the grammatical context was equivalent between sentence types, it seems most likely that the global rhythmic structure of the sentence was influencing the accuracy of their repetition. This surprising result strengthens the position of the Metrical Template Hypothesis, by providing data supporting the notion of a processing bias towards Sw structures at a sentence level in primary-school children.

Infants acquire a preference for the dominant Sw structure of English early on (Jusczyk, Cutler, & Redanz, 1993). This ‘trochaic bias’ has also been noted in experiments using tones (e.g.Abecasis, Brochard, Granot, & Drake, 2005; Brochard, Abecasis, Potter, Ragot, & Drake, 2003). Gerken (1994a, 1994b, 1996) argued for the supremacy of the Sw structure in the spontaneous speech of 2-year-old children. She found that children’s inclusion or omission of object articles was dependent on the potential for the resulting sentence to create Sw metrical feet and argued that the drive towards Sw structures overrode the syntactic constraints of the target phrasal structure. Trochaic bias was also evident here. If we consider that previous existing rhythmic templates are part of the linguistic knowledge drawn upon to complete a task, then a strong Sw template would provide a more robust template to scaffold subsequent repetition. Jusczyk, Houston, & Newsome (1999) found that the ability to segment wS structures was later developing in infants, suggesting that such a template is potentially less robust for English-learning infants. This could be particularly the case for children with DLD, who may have more difficulties in incorporating less dominant structures into their internal repertoire of possible templates.

Of particular interest was the relative nature of the influence of stress, such that the relative accuracy of repetition of strong and weak syllables was internally consistent within a word or phrase, but not across sentence-spans. Stress therefore appears to be a relative, rather than absolute, construct. We shall return to the implications of this in the General Discussion in Chapter 10.

We had hypothesised that provision of a regular sentence rhythm would support the laying down of the rhythmic template of the sentence and hence support repetition, however there was no significant effect of regularity on accuracy at sentence or syllable level. If we look to theories of temporal expectancy and neural oscillations, we can see a possible explanation for this lack of effect. In creating our stimuli, we varied the rhythm at the syllable level, however the stressed syllable beat remained constant. Stressed syllables (no matter the number of intervening unstressed syllables) always fell at consistent 750ms intervals. If it is the stressed syllables which direct temporal expectancies and attentional resources to moments in time, then there was in fact no difference between the timing of the regular and variable sentence rhythms in the stimuli at the stressed syllable level.

We had also hypothesised that providing an entraining rhythm would support repetition, whereas in fact it did not. There are two possible explanations for this. It could be that because the sentences were deliberately rhythmically produced, this in itself provided a sufficiently robust entrainment, rendering the additional cue redundant. Alternatively, it could be that our entraining rhythm provided insufficient information to be helpful at this larger, sentence level. Our results demonstrate that the repetition of rhythmic structures was determined by phrase-consistent levels of accuracy. Many accounts of prosodic phonology (e.g. Nespor & Vogel, 1986) suggest higher levels of prominence, beyond stressed and unstressed, which frequently coincide with phrasal and clausal level structures. By attempting to provide a simplified version of rhythm, focusing only on the lower two levels of prosodic structure, it could be that we inadvertently removed some of the most significant rhythmic information relevant for higher-level syntactic structures such as phrases and clauses. It could be that a more complex entraining rhythm which mirrored the overall pattern on several layers would have proved more useful for this particular task.

We would not, however, wish to contend that rhythm was the only influence upon repetition. There were also clear influences of grammar with the complexity of target structures also reflected in the children's accuracy patterns. In this task, the children found the embedded sentence structure particularly challenging. This structure was deliberately chosen as it is the final sentence-type included in the TROG and contains a particularly complex syntactic structure. When presented with these sentence types in the TROG assessment, very few of the children (four AMC, no DLD or YLC) passed the block, demonstrating that most of the children were unable to understand this structure accurately. This led to two forms of typical error pattern. Some children attempted to create a second 'tag' phrase e.g. 'the table the computer is under – it's blue'. Others appeared to attempt to create a prepositional phrase structure from the target preposition e.g. 'the table the computer is under the blue'. Here, they appear to have retrieved the preposition 'under' along with its usual surface syntactic form - a prepositional phrase- and then tried to embed the retrieved lexical item (here 'blue') within that structure. Here, the long-term grammatical framework which has elsewhere supported accurate repetition, is supporting the wrong conclusion, leading to higher error rates for this structure. Polišenská, Chiat & Roy (2015) found the grammaticality of target sentences to be a significant linguistic factor in sentence repetition in typically developing 4-5 year olds. Poor accuracy in repeating this sentence type would corroborate that finding by suggesting that these embedded clauses were unfamiliar to the children and so effectively violated grammaticality, since they contained a syntactic structure that was unknown.

Conversely, Polišenská et al. (2015) also found that familiar function word frames for a sentence were of more benefit than familiar content words. This may explain some of the findings in the thematic analysis. We found that Determiner accuracy was generally quite high, whilst accuracy of Auxiliaries and Copulas appeared to vary considerably. It could be that when the function words occurred within a familiar grammatical context (e.g. determiner + noun to create a noun phrase), the children were better able to repeat them than when the grammatical context was less familiar. Furthermore, the familiar phrasal structure may have created a stable syntactic template in the redintegration process, into which the content words could be slotted. Conversely, when the grammatical framework was potentially less secure due to its more advanced nature (e.g. the pluperfect auxiliary 'had' in the complex sentences), accuracy levels were poorer. Consideration of a 'Function Word Framework' may also explain in part the lack of effect found for prosodic context. The work which inspired this aspect of our task (Gerken, McGregor & Leonard) investigated function word production in very young children, for whom function words could be supposed to be a still developing aspect of their linguistic competence. Here, the children were much older, and even those with DLD could be reasonably supposed to be familiar with the 'determiner + noun' structure. Because the children could successfully draw on well-established grammatical knowledge to create this structure, this may have meant the prosodic context effect for these particular words was redundant. A separate consideration of the effect of prosodic context specifically for more grammatically challenging function words (e.g. 'had') may have yielded a different result. Familiarity of syntactic structure also therefore appears to play a role in successful sentence repetition.

Finally, we considered the relationship of sentence repetition with acoustic sensitivity. We hypothesised that poor acoustic sensitivity to aspects of rhythm would relate to sentence repetition performance. This was what we found in our correlation analysis, with sensitivity to Frequency, Duration and Rise Time (all contributors to percepts of rhythm and stress) related to task performance at Sentence, Word and Syllable level. The further regressions indicated that Frequency continued to provide significant amounts of unique variance above the influence of age and NVIQ at both Word and Syllable level, whilst Duration also provided significant unique variance at Syllable level. Sentence Repetition accuracy therefore improves with age and with increasing NVIQ, perhaps reflecting the complexity of the task, however acoustic sensitivity also plays an additional role in task success beyond the part played by these factors.

Rhythm has therefore been shown to play a role in sentence repetition, with the rhythmic structure of the target at syllable, word, phrase and sentence level influencing the accuracy of repetition. Stressed syllables appear to occupy a privileged place in this process. Individual repetition of stressed syllables is more accurate than unstressed, while unstressed syllables occurring together with stressed syllables as part of a target word are more accurately repeated by typically-developing children than those unstressed syllables which occur in isolation. The presence of a stressed syllable within a target structure therefore appears to be facilitating accurate repetition. The influence of rhythmic patterning at word-level was also evident, with the accuracy for individual syllables within words reflecting the word's internal rhythmic structure. Rhythm also influenced repetition at a global, sentence level, whereby sentences rich in Sw structures were more accurately repeated than those in which nonSw structures predominated, suggesting an overall processing advantage for Sw structures. Crucially, the privileging of stressed syllables occurred at a word or phrasal level, such that accuracy patterns were internally consistent, but with absolute accuracy levels varying between targets of a similar structure, even within the same sentence.

9 Experimental Task 5: Metrical Stress

9.1 Theoretical Basis of the Task

The final aspect of language and rhythm that we considered of interest was the potentially complementary nature of prosodic and syntactic structure operating across phrase and sentence level units.

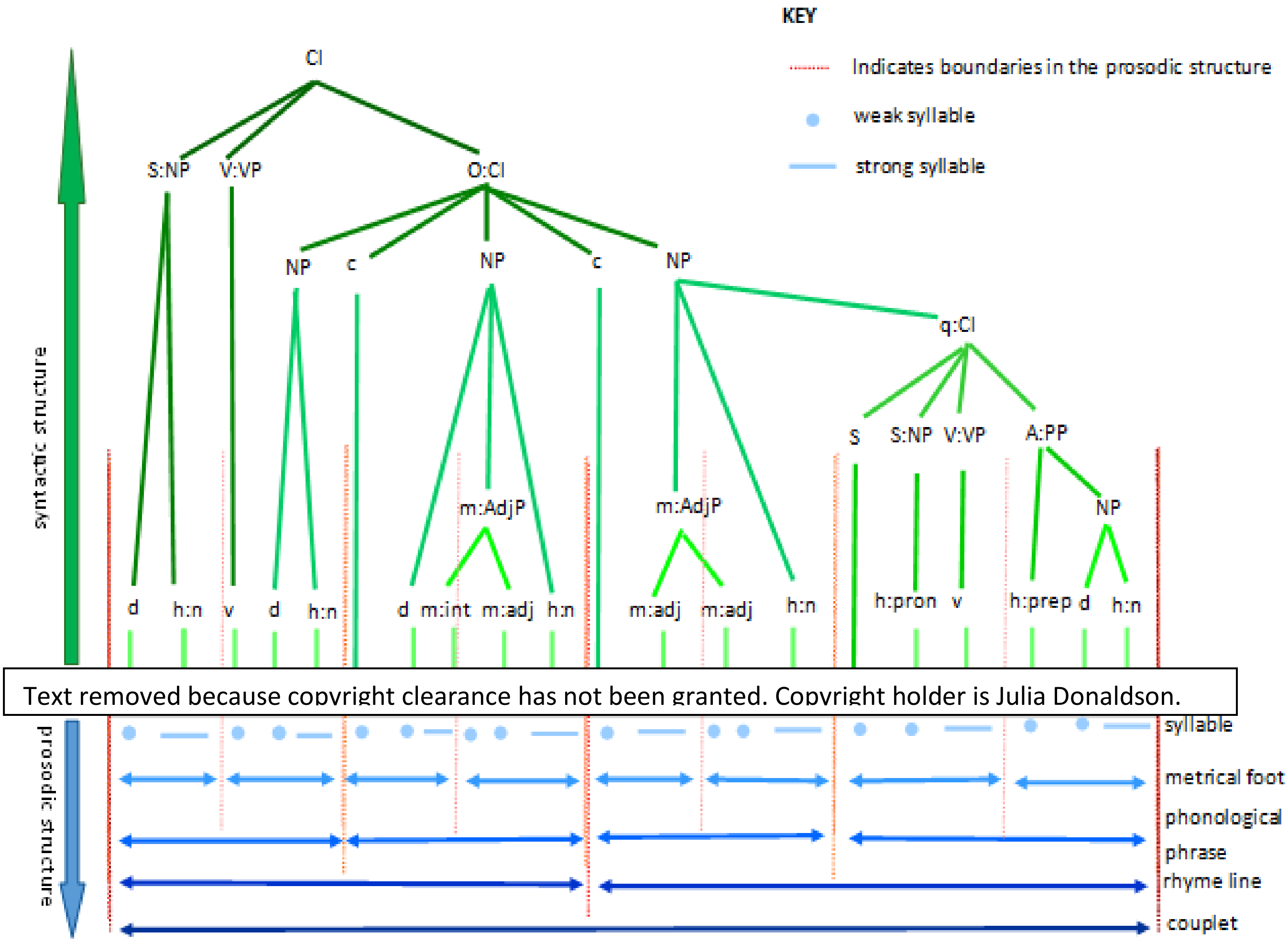
In linguistic accounts of prosodic structure theory, the prosodic structure of a sentence is thought to occur separately from that sentence's morpho-syntactic structure, with its own hierarchical arrangement (Selkirk, 1996). Nonetheless, the prosodic and syntactic structure of sentences often coincide, particularly in terms of the major groupings (Price et al., 1991) and are thought to have a more stable correlational relationship at the units of the phonological word and phonological phrase (Dresher, 1996). The existence of a prosodic-syntactic relationship, even if it is not a perfect match, is further strengthened by the argument that prosody and rhythm can contribute to the discovery of syntax by infants (Hirsh-Pasek et al., 1987; Jusczyk et al., 1992; Kemler Nelson, Hirsh-Pasek, Jusczyk, & Cassidy, 1989) – prosody and syntax clearly do not operate entirely independently of each other. If children with DLD have particular difficulty in using rhythm and prosody to discover syntax as we have argued above, then it could be that this difficulty is exacerbated when the relationship between prosodic and syntactic structures is more fluid, as linguists have argued is the case for natural language. Forms of linguistic expression in which the prosodic-syntactic relationship is particularly strong and consistent might, in contrast, be more supportive in helping children to discover the prosody-syntax relationship and therefore could be a useful means of targeting this area of impairment.

If we accept the view that there is an imperfect relationship between prosodic and syntactic structure in natural language, there is nonetheless a core part of children's typical language exposure in which the two levels of analysis are more tightly integrated, and that is in the realm of children's oral and textual culture. These aspects of children's linguistic life draw heavily on rhythm to structure the language – phenomena such as nursery rhymes and clapping games depend upon the integration of repetitive language and repetitive rhythm. A further aspect of a child's typical linguistic environment that draws on these features is in literature which relies upon rhythm and rhyme. Oeuvres such as those of Dr Seuss and the recent Children's Laureate, Julia Donaldson, build upon the playfulness of oral rhymes and games and their writing is characterised by a strong, repetitive framework of rhythm and rhyme. We hypothesised that the predominance of rhythm and rhyme in these texts may serve a scaffolding function in developing children’s awareness of prosodic-syntactic units.

The reasons for making such a hypothesis lie in the tight integration of prosody and syntax created by the rhythmic format and hence the richness of structural cues that it contains. Here, the child is exposed to cues at multiple hierarchical layers drawing their attention simultaneously to the phonological, prosodic and syntactic structure of the language. The property of rhyme emphasises the phonological structure of words by drawing attention to the onset-rime division, whilst also providing a guide to linguistic structure since each rhyme occurs at the end of a syntactic unit (be that clause or phrase). The overarching metrical structure also draws attention to the rhyme as a boundary point, since it occurs at regular intervals every four metrical feet. Within that metrical structure, there are further subdivisions into pairs of metrical feet, each of which also generally represents a complete syntactic unit. The metrical structure is therefore not an arbitrary form superimposed on the syntax of the text, but the two structures form a rich and highly integrated input which serves to highlight and reinforce the rhythmic and syntactic properties of language.

In order to illustrate this (Figure 9-1) we have taken the opening sentence of a popular children’s book (*Room on the Broom – Donaldson (2002))* and marked out the major syntactic structures (above the text in green) and the prosodic structures (below the text in blue). We can see that the major groupings in the syntactic structure are mirrored by major prosodic boundaries (the dashed red lines) in the prosodic structure. The prosodic boundaries are nested such that the larger the prosodic-syntactic unit, the greater the overlap of boundary cues, such that the end of each rhyme line represents the combined boundary of four different levels of metrical analysis as well as the boundary of a major syntactic unit.

The prosodic structure is built around the stressed syllables, which serve to demarcate the end of a metrical foot (predominantly anapaest i.e. wwS). The symmetry is not faultless, as can be seen from ‘a very tall hat’, in which the lexical word ‘very’ crosses the boundary of the metrical foot; however for the majority of the couplet there is a strong coincidence of prosodic and syntactic boundaries.



Text removed because copyright clearance has not been granted. Copyright holder is Julia Donaldson.

Figure 9-1 Diagram to illustrate the syntactic and prosodic structure of a line from Room on the Broom. Abbreviations: d-determiner; h:n -noun, head of noun phrase; v-verb; m:int-modifier:intensifier; m:adj-modifier:adjective; h:pron-pronoun, head of noun phrase; h:prep-preposition, head of prepositional phrase; c-conjunction, q-qualifier; Cl-clause; S-subject; O-object; A-adverb; NP-noun phrase; VP-verb phrase; PP-prepositional phrase

Given this level of dovetailing between the prosodic and syntactic structures, we were interested to find out to what extent the children with DLD were able to integrate these two systems of representation. If the children's ability to draw upon rhythmic patterning as a structure supporting syntax is diminished compared with typically developing children, then it could be that input which serves to emphasise these patterns could lead to more robust representations.

9.2 Developing the Task

The aim of the metrical stress task was to investigate whether children are sensitive to the coincident boundaries of prosodic and syntactic units as exemplified in the rhythmic texts which form a central part of children's literature.

Several rhythmic children's books were short-listed and then analysed for the regularity of the rhythmic and metrical structure. The one that best fitted the criteria was *Room on the Broom* (Donaldson, 2002). Each line of the book was then analysed according to its internal rhythmic pattern.

9.2.1 Characteristics of 'Room on the Broom'

The book as a whole is written in anapaestic tetrameter, with wwS as the dominant foot pattern.

The text is structured into rhyming couplets e.g.

Text removed because copyright clearance has not been granted. Copyright holder is Julia Donaldson.

Each line ending in a rhyme contains four stressed syllables:

Text removed because copyright clearance has not been granted. Copyright holder is Julia Donaldson.

Each line ending in a rhyme is composed of two syntactic units, each of which contains two stressed syllables (i.e. two metrical feet):

Text removed because copyright clearance has not been granted. Copyright holder is Julia Donaldson.

There is therefore a very clear and regular correspondence between metrical and syntactic units which is repeated throughout the text.

From the analysis of the whole text, ten couplets were chosen to form the stimuli set. Five couplets had the regular pattern:

SwwS; wwSwwS
wSwwS; wwSwwS

Text removed because copyright clearance has not been granted. Copyright holder is Julia Donaldson.

The other five couplets had the regular pattern:

wSwwS; wwSwwS
wSwwS; wwSwwS

Text removed because copyright clearance has not been granted. Copyright holder is Julia Donaldson.

9.2.2 Devising the stimuli

To investigate whether metrical groupings influence detection of syntactic-prosodic units, three conditions were created. A pause was created in the spoken recordings of the couplets according to three different criteria:

Metrical-Coincident (Met-Co) A pause at the end of every two metrical feet. This pause therefore coincided with natural prosodic and syntactic groupings.

Example:

Text removed because copyright clearance has not been granted. Copyright holder is Julia Donaldson.

wSwwS
wwSwwS
wSwwSw
wSwwS

Metrical – Non-Coincident (Met-NonCo): A pause which created regular metrical groupings but which did not coincide with syntactic boundaries.

Example:

Text removed because copyright clearance has not been granted. Copyright holder is Julia Donaldson.

wSw
wSwwSw
wSwSw
wSwwSw
wS

Non-Metrical – Non-Coincident (NonMet-NonCo): A pause which created irregular rhythmic groupings and also did not coincide with syntactic boundaries.

Example:

Text removed because copyright clearance has not been granted. Copyright holder is Julia Donaldson.	w SwwSwwSww SwSww SwwSwwS

It should be noted that the *syntax* in each version remains identical, only the prosodic grouping is altered. Accurate judgement is therefore not based on syntactic knowledge per se, but on how prosody and syntax typically interact.

9.2.3 Predictions

We predicted that if children have a robust representation of how prosody and syntax typically interact, they would be able to readily identify when this typical grouping was disrupted (Met-NonCo and NonMet-NonCo Conditions). We predicted this would be the case for AMC children.

Conversely, we predicted that children with DLD would have less well-established schema for prosodic-syntactic interaction and so would make more identification errors.

If children were sensitive to metre, but not how it integrates with syntax, then we would expect lower accuracy for the Met-NonCo condition. In contrast, if children were insensitive to both prosody and syntax, then we would expect no difference in accuracy between all three conditions.

9.2.4 Creating the Stimuli

To record the stimuli, a regular beat was induced in the speaker using a priming metronome beat in one ear (inaudible on the recording) set to 80 beats per minute (equivalent to 750ms between beats). The stimulus was then spoken so as to align the stressed syllables with the beats at 750ms intervals. The precision of this timing was then verified and adjusted as necessary using Audacity software.

The inserted pause was equivalent to the insertion of one silent stressed syllable interval, such that there was 1500ms between the preceding and following stressed syllable.

A preceding entraining beat was created using audiofiles created by Matlab software and inserted into the recording using Audacity. The entraining beat consisted of the equivalent of eight metrical feet, thereby paralleling the metrical structure of the spoken stimulus.

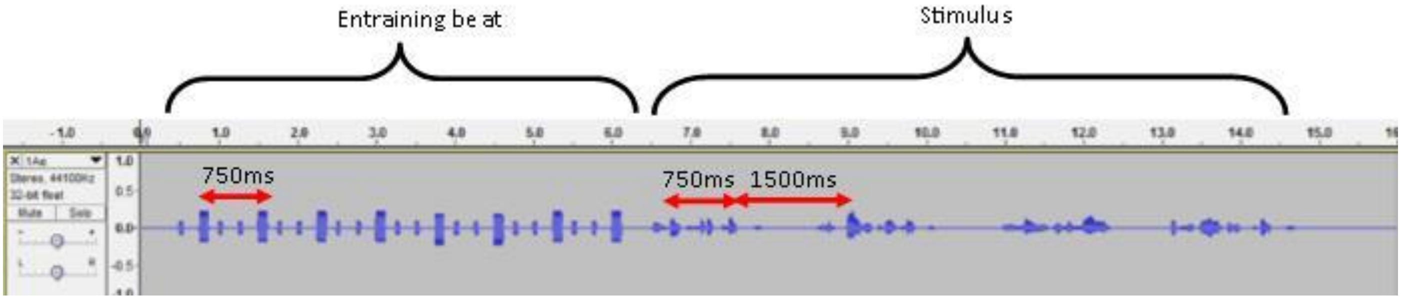


Figure 9-2 Example Soundwave for Condition Met-Co:

Text removed because copyright clearance has not been granted. Copyright holder is Julia Donaldson.

Each couplet was recorded in three different versions: Met-Co, Met-NonCo, and NonMet-NonCo. The couplets were then arranged in three blocks of ten, with each block containing a counterbalanced mix of all three conditions (e.g. four Met-Co, four Met-NonCo and three NonMet-NonCo). Each couplet occurred only once in each block, and the order of couplets was fixed was across blocks. Each block was listened to in a separate session, with the order of presentation of blocks across the three sessions randomised across participants. Each child ultimately listened to each block and therefore recorded scores for all three versions of each couplet.

9.2.5 Procedure

In the first testing session, the experimenter read the entire storybook to the child so that each child was familiar with the entire text. Each experimental presentation of the task then occurred as part of the wider task battery in subsequent sessions.

The task was contextualised by talking about how when reading out loud it was important to take a breath in a ‘sensible place, where it fits with the words’ because otherwise ‘it…sounds interrupted … like…this.’ It was then explained that they were going to hear someone reading the words from ‘Room on the Broom’ but that sometimes the reader would take a breath in a ‘funny place, where it sounds wrong, like it doesn’t fit’. They were then instructed that if the breath sounded like a sensible place where it fitted with the words, they should press the ‘tick’. If it sounded wrong, or as though it was interrupted, they should press the ‘cross’.

The corresponding picture from the book was displayed on screen whilst each stimulus played.

Prior to the presentation of each block, three practice trials were completed, during which feedback was given to ensure the children understood the task. The ten experimental stimuli of that block were then presented.

9.3 Results

9.3.1 Accuracy Scores

The children's scores were calculated according to number of responses correct (i.e. identifying Stimulus Type Met-Co as correct with a √ press, and Met-NonCo and NonMet-NonCo as incorrect with a x press). The maximum score was therefore 30, with a maximum score of 10 for each Stimulus Type.

A one-way ANOVA (DV-Score) revealed a significant group difference in overall accuracy, *F*(2, 51) = 11.986, *p* = .000. Post-hoc comparisons (Bonferroni) showed that the AMC group were more accurate than the DLD (*p* = .007) and YLC groups (*p* = .000), whilst there was no group difference between the DLD and YLC groups[24].

Table 9-1 Results of one-way ANOVAs for Score

Score (Overall - Max 30; Individual - Max 10)	**AMC Mean (SD)**	**DLD Mean (SD)**	**YLC Mean (SD)**	***df***	***F***	***p***
Overall Score[a,b]	26.14 (3.692)	20.33 (5.263)	18.81 (5.980)	2, 51	11.986	.000
Met-Co	7.57 (2.908)	7.25 (3.166)	7.71 (2.217)	2, 51	.112	.895
Met-NonCo[a,b]	9.05 (1.244)	6.08 (2.999)	5.67 (3.454)	2, 22.475[c]	12.603	.000
NonMet-NonCo[a,b]	9.33 (1.278)	6.75 (3.251)	5.43 (3.187)	2, 22.513[c]	15.325	.000

Note: a = AMC > DLD; b = AMC > YLC; c) Welch's df and F used due to significant Levene's test

---

[24] Due to software errors, two children from each group unintentionally listened to the same block presentation twice. These children's scores were removed from the summary analysis. One AMC child had scores which were identified as outliers, so this child's score was also removed.

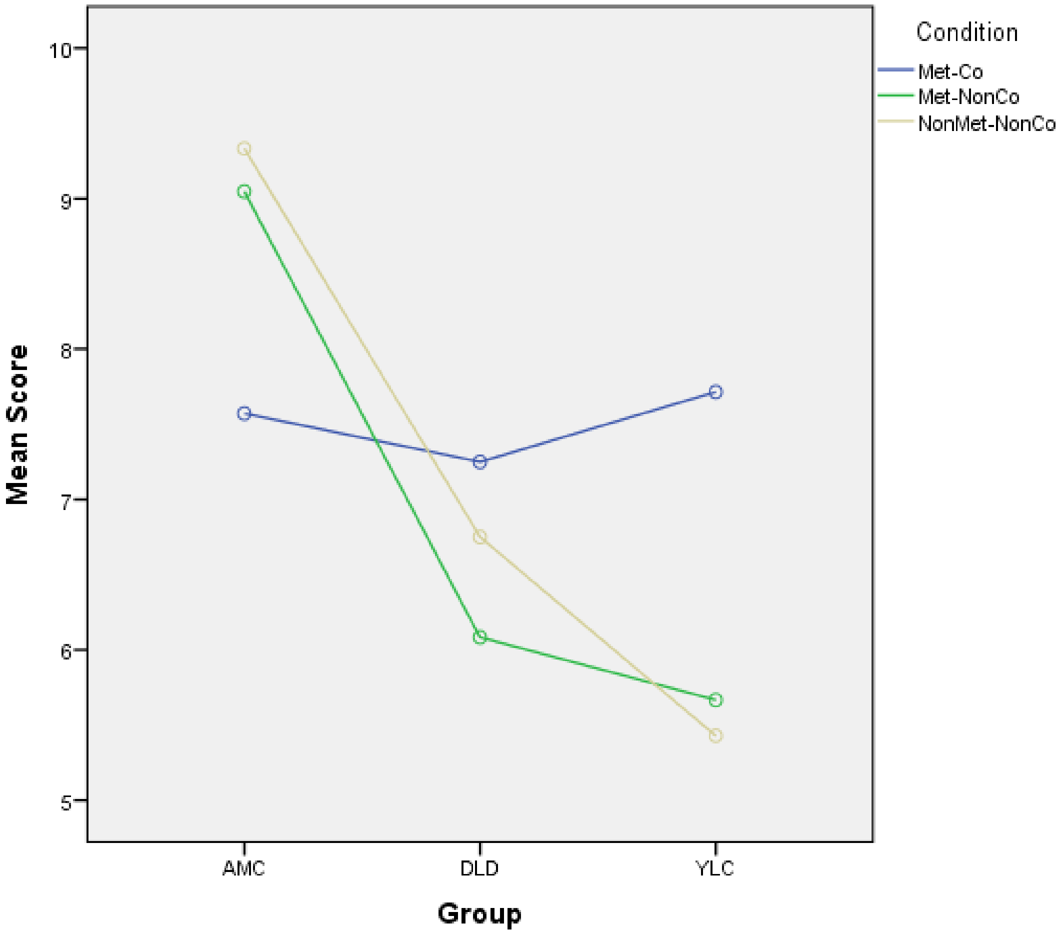


Figure 9-3 Graph of Mean Scores by Group for each Condition

Repeating the one-way ANOVA for each Condition showed that there was no effect of group for Met-Co (where metrical and syntactic structures were congruent), but a significant effect of group for Met-NonCo and NonMet-NonCo. In both cases, the AMC group were more accurate than the DLD group (*p* = .015; *p* = .05) and the YLC group (*p* = .001; *p* = .000) [Games-Howells comparisons]. There were no significant differences between DLD and YLC groups for any of the Conditions.

Children with DLD were therefore less accurate at identifying when the coincidence of prosodic-syntactic cues was disrupted than were typically developing AMC children. YLC children were also less accurate than the older AMC children.

Having established accuracy differences by group, we were interested in whether there was a significant difference in children's accuracy between the three Conditions, as this would help us understand whether children had difficulty in integrating the prosodic and syntactic structures. Children who have a well-developed schema for how prosody and syntax typically interact should have no difficulty in identifying the Met-NonCo and NonMet-NonCo Conditions as incorrect. Conversely, we predicted that children with DLD would have less established schema and so make more identification errors. If the children with DLD were sensitive to metrical structure but did not relate this to syntactic structures, then we would expect lower accuracy for Met-NonCo than NonMet-NonCo. If children were insensitive to both rhythm and its relationship to syntax, then we would expect no difference in performance between the three conditions.

In order to explore this, a series of one-way repeated-measures ANOVAS (Condition- Met-Co, Met-NonCo, NonMet-NonCo; DV-Score) were run for the whole dataset and then for each group individually (see Table 9-2).

Table 9-2 Results of repeated-measures ANOVA by Condition and Group

Score (Max 10)	**Met-Co**	**Met-NonCo Mean (SD)**	**NonMet-NonCo Mean (SD)**	***df***	***F***	***p***
AMC	7.57 (2.908)	9.05 (1.244)	9.33 (1.278)	1.072, 21.432[a]	5.106	.032
DLD	7.25 (3.166)	6.08 (2.999)	6.75 (3.251)	1.148, 12.625[a]	.413	.560
YLC	7.71 (2.217)	5.67 (3.454)	5.43 (3.187)	1.299, 25.988[a]	4.397	.037
All Groups	7.56 (2.675)	7.07 (3.083)	7.24 (3.120)	1.182, 62.650[a]	.442	.541

Note: a) Greenhouse-Geisser corrections for df and F due to significant Mauchly's test

There was a significant main effect of Condition for the AMC and YLC groups, although pairwise comparisons (Bonferroni) between each pair of conditions did not reach significance. There was no effect of Condition for the DLD group.

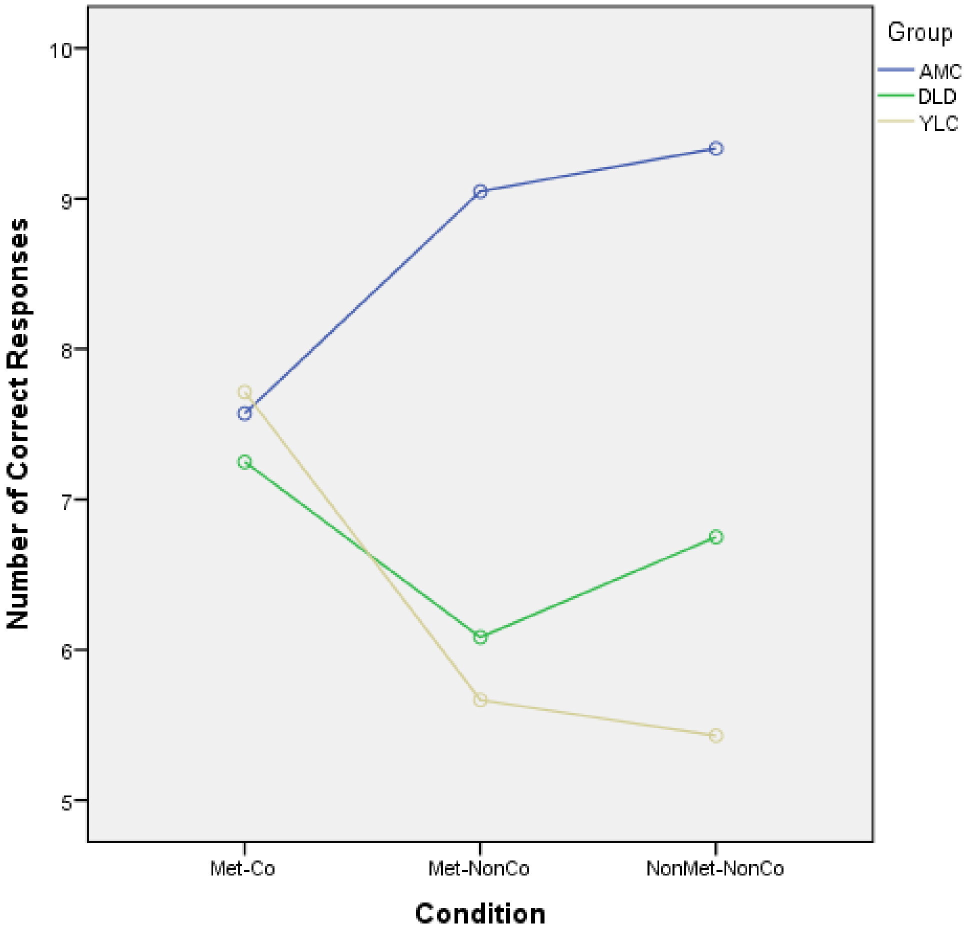


Figure 9-4 Graph showing the Mean Score for each Condition by Group

Figure 9-4 helps to illustrate that the direction of effect of Condition for the AMC and YLC groups was different, with the AMC group scoring more highly for the non-coincident Met-NonCo and NonMet-NonCo conditions than Met-Co, whilst the YLC group's scores were lower for the non-coincident stimuli than the coincident Met-Co type.

The AMC children scored highly for both non-coincident Conditions, indicating that they were reliably able to identify when rhythmic and syntactic structures did not coincide. In comparison, a series of one-sample t-tests (test value 5) revealed that the DLD and YLC groups did not score significantly different from chance for Met-NonCo or NonMet-NonCo conditions (DLD: Met-NonCo: $t(11) = 1.251$, $p = .237$; NonMet-NonCo($t(11) = 1.865$, $p = .089$; YLC: Met-NonCo: $t(20) = .884$, $p = .387$; NonMet-NonCo: $t(20) = .616$, $p = .545$. This suggests that the DLD and YLC groups were not able to reliably identify when the rhythmic boundaries did not coincide with syntactic boundaries.

We had predicted that children with DLD would have less developed processing of the integration between prosody and syntax than typically developing children, whereas TD children would be sensitive to the disruption of prosodic and syntax integration. This proved to be the case, for the older children with the AMC group correctly able to identify the Met-NonCo and NonMet-NonCo stimuli as ‘incorrect’. Conversely, the children with DLD did not reliably identify the non-coincident stimuli, nor did the YLC group. The children with DLD had a small dip in accuracy for Met-NonCo scores compared with Met-Co and NonMet-NonCo, a pattern consistent with being sensitive to rhythm (i.e. detecting the metrical structure) but not its relationship with syntax. However, there was no overall significant difference for the DLD group between the conditions and so a more conservative conclusion would be that the DLD group did not consistently process the metrical structure nor were able to relate the structure to syntactic units.

An unexpected pattern in the results was the relatively poor performance of the AMC group in the Met-Co condition. This means that this group were more likely to reject a correct rendition than to accept an incorrect one.

9.3.2 Reaction Times

Reaction Times (RTs) were measured for each response at a time resolution of 0.1ms and then a mean RT was calculated for each child for each Condition.

A repeated-measures 3 x 3 ANOVA (Group-[AMC, DLD, YLC] x Condition – [Met-Co, Met-NonCo, NonMet-NonCo) showed no significant effect of Condition on RT (*F*(2, 102) = .101, *p* = .904), and a marginally significant effect of Group (*F*(2, 51) = 3.112, *p* = .053. Pairwise comparisons (Bonferroni) showed that the YLC group were significantly slower to respond overall than the AMC group (*p* = .049), but there was no significant difference in RT between the AMC and DLD groups (*p* = .610).

There was also a significant Group*Condition interaction (*F*(4) = 5.229, *p* = .001) – see Figure 9-5.

Table 9-3 Results of one-way ANOVAs by Group for each Condition

Reaction Time **0.1ms units**	**AMC** **Mean (SD)**	**DLD** **Mean (SD)**	**YLC** **Mean (SD)**	***df***	***F***	***p***
Met-Co	475132.94 (42717.59)	472686.78 (74534.29)	504839.29 (74358.18)	2	1.477	.238
Met-NonCo	444700.69 (91677.26)	506524.35 (83676.8)	502674.99 (80292.64)	2	3.092	.054
NonMet-NonCo[a]	448344.47 (71947.07)	488480.9 (75708.31)	524136.54 (83377.62)	2	5.037	.01

Note: a) AMC < YLC;

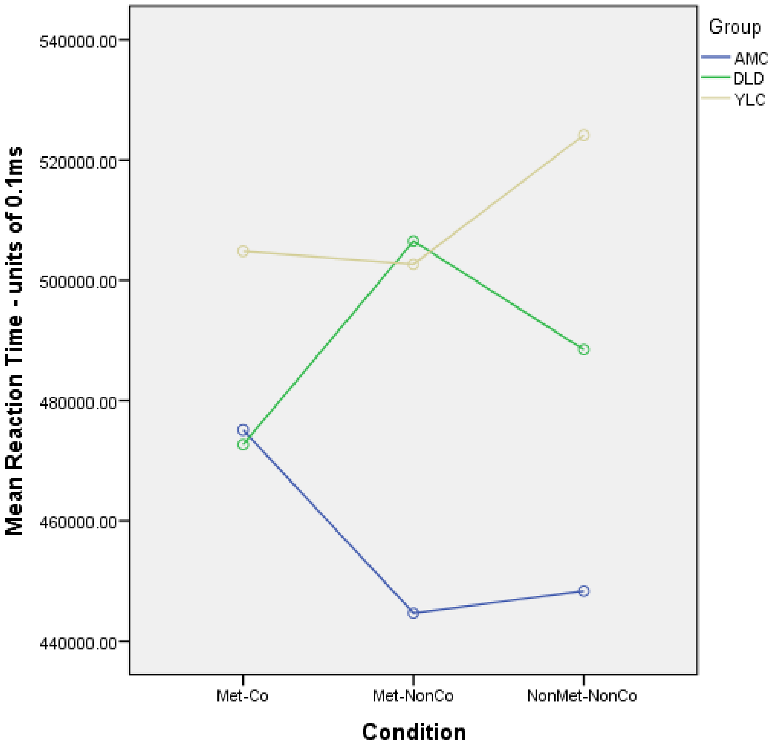


Figure 9-5 Graph of mean reaction times to each Condition by group

We can see from the graph of mean RTs the implication of the Group*Condition interaction, as each group shows a different response profile. The AMC group tend to be quicker to respond to the non-coincident stimuli (Met-NonCo, NonMet-NonCo) than the coincident (Met-Co) stimuli. The DLD response pattern differs from the AMC group, with Condition Met-NonCo resulting in their slowest responses, but little difference between types Met-Co and NonMet-NonCo. The slowest reactions for the YLC group were for NonMet-NonCo.

Table 9-4 Results of repeated-measures ANOVAs for Reaction Time by Group

Reaction Time 0.1ms units	**Met-Co Mean (SD)**	**Met-NonCo Mean (SD)**	**NonMet-NonCoMean (SD)**	***df***	***F***	***p***
AMC[a]	475097.889 (42704.323)	444422.536 (92030.43)	451399.2302 (67109.5167)	1.359, 27.185	4.623	.03
DLD[b]	472395.438 (77190.77096)	506524.35 (83676.8)	491110.3403 (77033.21619)	2, 22	4.156	.029
YLC	504839.29 (74358.18)	502016.5119 (80688.314)	524122.52 (83378.20)	2, 40	2.387	.105

Note: a) Met-Co > NonMet-NonCo; b) Met-NonCo> Met-Co

To understand the group profiles further, a series of one-way repeated-measures ANOVAs for Condition were run individually by group. There were significant effects of Condition for the AMC and DLD groups, but no significant effect for the YLC group.

The effects of Condition for the AMC and DLD groups were in opposite directions. For the AMC group, responses to the Met-Co (Metrical-Coincident) Condition were significantly slower than to the NonMet-NonCoCondition (*p* = .049). This means that it took the AMC group longer to decide that the coincident stimulus was correct than to decide that the non-coincident stimulus was incorrect. This result tallies with observations during testing – children often pressed the [x] button as soon as they heard the first non-coincident boundary, immediately confident that this presentation was 'wrong'. In contrast, in order to be sure that the coincident stimulus was correct, the stimulus had to be listened to in its entirety. This may explain this difference in response times for the AMC group.

In contrast, the DLD group were rarely observed to press the response buttons before the full stimulus was played, explaining their generally slower reaction times. We had predicted that the Met-NonCo stimuli might cause difficulties if the children were sensitive to the metrical rhythm, but were unsure of the metrical rhythm should interact with syntactic boundaries. We found no significant effect on accuracy for the Met-NonCo stimuli, but the DLD RTs for Met-NonCo condition were significantly slower than for Met-Co (*p* = .034). This provided support for the idea that integration of rhythmic and syntactic cues took longer to resolve in this conflicting condition for children with DLD.

9.4 Relationship with Acoustic Thresholds

We predicted that success in this task may be related to acoustic sensitivity to RiseTime, Frequency and Duration, with better acoustic sensitivity related to better task performance. In order to investigate this, we conducted correlations between the four acoustic threshold measures (Chapter 4) and score on the Metrical Stress task.

Table 9-5 Correlation co-efficients (Pearson one-tailed) for Acoustic Thresholds and Metrical Stress Score - AMC and DLD groups

	Duration	Frequency	Intensity	Task Score
Rise Time	.340*	.795***	-.227	-.363*
Duration		.406**	.016	-.526**
Frequency			-.206	-.565***
Intensity				-.241

Note: *$p < .05$; **$p < .01$; ***$p < .001$

There was a significant correlation between performance on the Metrical Stress task and sensitivity to Rise Time ($p = .019$), Duration ($p = .001$) and Frequency ($p = .000$). The inverse relationship indicates that the smaller the discrimination threshold (i.e. better the acoustic sensitivity), the better the task performance.

A set of regression equations was conducted to explore the unique variance in Metrical Stress score accounted for by the predictors of Age, NVIQ and each AT measure in turn. Step 1 was always entered as Age (months), Step 2 as NVIQ and Step 3 as the AT measure. The overall accuracy score was entered as the dependent variable.

Table 9-6 Results of Regressions exploring the unique variance in Metrical Stress Score accounted for by Age (months), NVIQ and AT measures (AMC and DLD groups)

	b	SEb	β	*t*	*p*	ΔR^2	*p*
Model 1							
Age	.016	.056	.046	.291	.773	.067	.151
NVIQ	1.315	.413	.528	3.187	.004	.296	.001
Rise Time	-.010	.010	-.160	-1.002	.325	.022	.325
Model 2							
Age	.009	.054	.027	.174	.863	.067	.151
NVIQ	1.070	.431	.429	2.482	.019	.296	.001
Duration	-.031	.017	-.305	-1.795	.083	.066	.083
Model 3							
Age	-.009	.047	-.027	-.182	.857	.040	.278
NVIQ	1.033	.379	.465	2.729	.011	.353	.000
Frequency	-2.537	1.295	-.324	-1.960	.060	.076	.060
Model 4							
Age	.026	.058	.075	.459	.650	.071	.148
NVIQ	1.446	.407	.580	3.551	.001	.304	.001
Intensity	-.117	.535	-.034	-.219	.829	.001	.829

Age was not a significant predictor of performance in this task (*p* range = .148 - .278), however NVIQ contributed significant amounts of unique variance throughout the models (range 29.6% - 35.3%, *p* range = .000 - .001). The greatest variance accounted for by the AT tasks was for Frequency (7.6%) and for Duration (6.6%) although these changes did not reach significance (*p*= .06, .083 respectively). Neither Rise Time nor Intensity made significant contributions to Score in this task (*p* = .325, .829 respectively).

The significant correlations between Rise Time, Duration and Frequency sensitivity and Metrical Stress score may therefore have been partly mediated by NVIQ, with the acoustic cues of Frequency and Duration providing smaller (non-significant) additional contributions to the variance in score.

9.5 Summary of Findings from the Metrical Stress task

Children with DLD were less able to detect when prosodic-syntactic boundaries were non-coincident than were typically developing children. Further, they were more likely to accept the non-coincident stimuli as correct than were AMC children, although the accuracy of responses between individual Conditions did not differ. The profile of Reaction Times for DLD children also differed from the AMC group, as they had significantly slower reactions for the Met-NonCo Condition, whilst AMC children were slowest for Met-Co stimuli[25].

9.6 Discussion

This task set out to investigate the influence of metrical groupings on children's capacity to recognise when typical prosodic-syntactic groupings were interrupted. We considered that the metrical regularities of children's rhythmic texts might serve to emphasise the syntactic structures of language through the coincidence of prosodic and syntactic boundaries. If children have robust knowledge of how prosody and syntax interact, then any violation of these coinciding units should be readily identified. Alternatively, if they are able to detect rhythmic patterns but are unable to relate these to the overall prosodic-syntactic structure, then the condition in which there is a metrical rhythmic structure which does not coincide with the syntax (Condition Met-NonCo) should prove more problematic than when there is no consistent metrical pattern (Condition NonMet-NonCo). If the children are insensitive to both prosody and its relationship with syntax, then there should be no difference between the three different conditions.

All three participant groups had a similar success rate with Met-Co stimuli (correct prosodic-syntactic groupings). However, the AMC group were more successful in detecting the non-coincidence of metre and syntax in the Met-NonCo and NonMet-NonCo conditions, whilst DLD and YLC groups failed to reject the non-coincident stimuli. It seems therefore that it was the AMC group who did surprisingly poorly in condition Met-Co (in the context of their other responses) rather than that the DLD and YLC groups did surprisingly well. There was no significant difference in accuracy between the three conditions for the DLD group, indicating that they found each condition equally challenging to interpret. The DLD group were therefore impaired compared to the AMC group at detecting violations of prosodic-syntactic units and their accuracy pattern was similar to that of the younger children. This is the result pattern that we would expect if the children were having difficulty processing language metre and its relationship with syntactic structures.

---

[25] Recall that all children completed Entrained versions of this task only. Accordingly, there are no comparisons of Unentrained and Entrained scores for this task.

For the DLD group, the metrical grouping of the stimulus did not significantly affect accuracy of performance. However, taking a broader look at the results across the dataset, we can draw conclusions about the influence of prosody and metre on children's responses. If the children's responses were influenced by metrical regularity per se (regardless of the syntax), we would have expected higher error rates for Condition Met-NonCo, with the children falsely responding to the regular metre. This was, in fact, the trend of results for the DLD group, but not significantly so. In contrast, if the children's responses were primarily motivated by the mismatch between prosody (regular or otherwise) and syntactic groupings, we would expect that Conditions Met-NonCo and NonMet-NonCo would have equally distributed scores and this is what was found for all three groups.

From the accuracy scores, the most significant impact of metrical grouping appears to be the attention that it draws to the prosodic-syntactic unit, rather than in its temporal regularity per se. Violations of this unit by disrupting the integration of prosody and syntax were readily identified by the typically developing children in our sample. For the younger children, they were less able to reject these forms, being more likely than the older children to accept them as correct. This suggests that younger children have less well-developed schema for how prosody and syntax typically interact than do the older children. There are also implications for the DLD children. They were also less able to reject the sequences in which the prosodic-syntactic units were violated and so also appear to have under-developed schema for the interaction of prosody and syntax. This suggests that they may not be processing all of the cues available to them in segmenting the speech stream into prosodic units and grammatical clauses. Note that the syntax itself was identical in all three stimuli sets, so the test is not one of grammatical structures per se, but of how these structures interact with prosodic units in typical speech.

The range of responses here sits interestingly between experiments with infants, which have shown that infants are sensitive discriminators of pauses inserted within clauses or at clause boundaries (Kemler Nelson, Hirsh-Pasek, Jusczyk, & Cassidy, 1989), and those with adults, who also judge sentences where pauses coincide with clause boundaries or phrasal boundaries to be more natural (Jusczyk et al., 1992). The older AMC children therefore were more adult-like in their responses, being able to judge the sentences that were not Met-Co as being unnatural. The question, however, is why the DLD and, particularly the YLC children have relatively poor accuracy if 9-month-old infants are sensitive to these boundaries? One explanation could be in task demands. In our experiment, the children were asked to explicitly decide which was the 'correct' version, and so this required a greater degree of metalinguistic awareness than the infants in Jusczyk's passive listening study. Our experiment also manipulated a range of phrasal and clausal structures rather than concentrating on just one form of boundary. This may have made the task more complex to carry out. In his discussion of Jusczyk's experiments, Eimas (1996) points out that pausing at phrasal boundaries is not a feature of naturally occurring speech, therefore arguing that phrasal boundary pauses may not be a cue for syntactic structure per se, but may serve to highlight other cues such as decreases in fundamental frequency. This notion of a combination of cues is an interesting one when we consider the possible scaffolding function of children's texts such as Room on the Broom. Here the prosodic boundaries indicated by metre, rhythm and rhyme combine to demarcate phrases as well as clauses, depending on the structure of the carrier sentence. It could therefore be that they are providing a richer, more structured input which serves to highlight the more subtle cues which may be available in more natural language. Indeed, Jusczyk et al. (1992) found that their stimuli had more robust acoustic indicators of phrasal units when they were obtained from storybook readings than from spontaneous speech, again pointing towards structured child-orientated input as having specific, beneficial prosodic indicators of syntax.

We found significant correlations between the accuracy score and the acoustic cues of Rise Time, Duration and Frequency. Further regressions indicated that Age was not a significant predictor of task accuracy whilst NVIQ contributed significant unique variance. This suggests that some of the significant correlations may have been mediated by NVIQ, with a lesser role for acoustic sensitivity itself in task success. Of the acoustic cues, the regressions found that Duration and Frequency contributed the greatest amount of additional (non-significant) unique variance in this task.

The Reaction Time data also provides valuable insight into the influences of metrical grouping, with each group showing a different RT profile influenced by Condition. If children were confident in using prosody to detect syntactic boundaries regardless of metre, we would expect swift responses to those violations (Met-NonCo and NonMet-NonCo conditions). Alternatively, if the children were able to detect the metrical grouping but could not readily relate that to the syntactic groupings, we would expect Met-NonCo stimuli to produce slower responses due to the conflicting information.

The AMC group were quicker to respond when presented with an incorrect stimulus (Met-NonCo and NonMet-NonCo) than Met-Co. This indicates that they were confident in detecting prosodic-syntactic boundary violations. This interpretation was borne out by observations during testing as many of the children responded as soon as they heard a pause which they considered incorrect, often without waiting to hear the whole recording. In contrast, it was necessary for them to hear the whole of the 'correct' recording to be sure that there were no errors, which would explain the slower RTs recorded for Met-Co stimuli for this group.

In contrast, we can see that the children with DLD did not make early detections of errors, almost always choosing to listen to the whole recording. Furthermore, it took them longer to decide on a response when there was a rhythmic metrical grouping (Condition Met-NonCo). This is the result that we would expect if they were having to decide between the divergent influences of metre and syntax. When the rhythm was non-metrical and non-coincident, there was no significant difference between their RTs for this and for the Met-Co stimuli. The RTs then reveal additional information to the accuracy scores. It seems that rather than being impervious to metrical structure (as the scores might indicate), the children are able to detect its presence (since it was able to exert a negative influence on their RT) but are unable to relate this reliably and easily to the syntactic structure it interacts with.

A possible interpretation of this effect can be found in the infant literature. The notion of metrical patterning serving a scaffolding function was suggested by Kemler Nelson et al. (1989). In their study, infants were able to detect non-coincident clause boundaries in infant-directed speech (IDS) but not in adult-directed speech (ADS). This led to the suggestion that the prosodic characteristics of IDS might serve to bind the elements of the clause together to create a unified perceptual whole. Metrical patterning provides a possible means for achieving this. In the experiment presented here, metrical patterning, as exemplified by children's literature rather than IDS, could be interpreted as having a unifying effect, such that for the DLD children, they puzzled longer over the non-coincident metrical stimulus than for the non-coincident non-metrical stimulus. In this sense, the metrical patterning may have been demarcating a unified perceptual unit, but across an erroneous segment of the sentence.

The discrepancy between infant sensitivity to prosody in IDS and ADS is also interesting when we look at the results here. The children with DLD appear able to respond to metrical structure when it is tightly structured, strong and reliable as it was in the readings in this experiment. All the rhythmic cues here were therefore exaggerated in comparison to natural speech. This is not dissimilar to the exaggeration of acoustic cues that we find in IDS, but which are less reliably found in ADS. Studies of the acoustic characteristics of IDS have found that the rhythmic focus rapidly shifts as the infant ages (Leong, Kalashnikova, Burnham, & Goswami, 2014). It could be therefore that for children with DLD, they require a longer period of structured prosodic input in order to develop their sensitivity to the meaning of the units. If children with DLD are less efficient at discovering these, and IDS changes rapidly with the age of the child, it could be that they therefore end up 'missing out' on this crucial early aspect of language acquisition as the incoming signal 'moves on' before their system is ready to cope with a less structured and salient input. Such a scenario would have significant implications for language development. Morgan & Saffran (1995) argued for regarding prosody as a kind of parameter-setting device, providing a rough categorisation of the input into smaller units, thereby constraining the amount of input which can then be subject to further analysis, such as statistical learning. If this is the case, then sensitivity to prosodic units is a powerful tool in the process of discovering grammatical units. In this view of language acquisition, a poorer sensitivity to prosody would mean that constraining parameters fail to be set in chunking the input stream, and so subsequent analysis would be carried out across much greater distributions of input, resulting in a far more inefficient and unwieldy task. This in turn would lead to difficulty in segmenting language into grammatical units such as clauses and phrases, with knock-on implications for acquiring smaller-grained aspects of morphology – exactly the kinds of linguistic difficulties which characterise DLD.

The results here, though, demonstrate that when the rhythmic signal is robust, reliable and structured, it can exert an influence on the ability of children with DLD to perform a language-based task, which suggests that exposure to strong, rhythmic language may still have a beneficial role for these children's language development. If this is the case, then this finding may have clinical implications. If the children are able to detect strongly marked rhythmic patterns, but not integrate them into prosodic-syntactic units, then this could be a valuable area of intervention. If prosodic cues can enable us to more readily parse the speech stream, then teaching children to listen for these cues may increase their ability to integrate these two sources of information. In this respect, children's capacity to derive grammatical structure from prosodically-driven mechanisms could be increased. Intervention using rhythmic children's books to highlight this congruence of prosody and syntax could theoretically be of great value in scaffolding development of this metalinguistic skill.

10 General Discussion

The study comprised five experimental tasks, each of which investigated a different aspect of language processing which we hypothesised would be influenced by rhythmic processing and therefore affect the performance of the children with DLD. Throughout the tasks, when rhythm and stress patterns were manipulated, this influenced the nature of children's responses. This set of results supports a role for rhythm and stress in language processing across linguistic levels.

10.1 Stress, Rhythm and Language Processing

For syllables, stressed syllables in repetition tasks (Nonword Repetition – NWR, Sentence Repetition – SR) were consistently more accurate than unstressed (All groups), suggesting a privileged status for stressed syllables in phonological memory. At a word level, words containing stressed syllables were more accurately repeated than unstressed monosyllabic words (SR-AMC group only). Together with the NWR results privileging stress, this supports a view of language processing in which the stressed syllable is the 'anchor' around which subsequent specification of lexical rhythmic and phonological structure is built. Infant studies argue for rhythm and stress as the initial representational level (e.g. Mehler et al., 1996) and we can see a parallel function for stress here in older primary-school children. Further evidence for the primacy of stress in representations comes from the NWR task in which stressed syllable accuracy rates were relatively constant across wordlengths, whilst unstressed accuracy rates declined for longer words. As the memory challenge increased, it was the unstressed syllables which suffered, whilst the stressed syllable provided the anchor supporting pre- and post-stress unstressed syllable repetition.

The patterning of stress at a word level also affected children's responses. From NWR, we discovered that individual stress patterns of words affect accuracy rates, particularly when the wordlength appeared to begin to challenge memory capacity. Interestingly, this did not correspond to frequency rates of stress patterns in English, suggesting that prior word-level templates are not as important a factor in NWR as we had supposed. An alternative explanation may be the individual acoustic characteristics of the word waveform itself.

Despite the lack of evidence for use of prior rhythmic templates being invoked in the NWR task, elsewhere in the results, there was evidence for internalised rhythmic templates facilitating language processing as proposed by Gerken (1996) and Cutler (1996). Developmentally, infants demonstrate a Sw processing bias (Jusczyk et al., 1999) reflecting the dominant rhythmic patterns of English (Cutler & Carter, 1987). In our data from older children, we also saw evidence for a Sw processing bias. Carrierphrases with a Sw structure resulted in faster responses (Lexical Stress - LS) whilst sentences rich in Sw structures had higher overall repetition accuracy levels than those in which nonSw structures predominated (SR). This corresponds to evidence from other fields in which adults display a binary Sw preference (Abecasis et al., 2005; Brochard et al., 2003). It seems that the Sw preference set up early in development continues to be a feature of language processing beyond the infant stage.

AMC children were able to readily identify disruptions of the typical interface between rhythm and syntax (Metrical Stress- MS). This suggests that whilst rhythm and other language structures may have some level of autonomous structuring (Selkirk, 1996), there must nonetheless be well-established expectations by late-primary age of how the two levels typically interact, suggesting an integration rather than dissociation of linguistic and rhythmic domains.

Language rhythm has therefore been demonstrated to form part of language processing at syllable, word, phrase and clausal levels, with stressed syllables emerging as a core element of linguistic representation.

10.2 Entrainment

A central feature of the design of our tasks was investigating the effect of providing an entraining rhythm. We had expected, based on theories of dynamic attending, temporal expectancy and oscillatory phase-resetting, that the addition of an entraining rhythm would facilitate both accuracy and efficiency of task response. This was not the general picture we obtained, however, as for the majority of our analyses there was no significant difference in results between Entrained and Unentrained versions of tasks. We must therefore conclude that in general, listening to an entraining rhythm did not facilitate task completion.

In considering possible reasons for the absence of this predicted effect, the answer could lie either in the nature of the target stimuli, or in the nature of the entrainment rhythm itself.

Regarding the target stimuli, these were deliberately rhythmic, recorded to coincide with a metronome beat. It could be that this increased rhythmicity may have induced a strong enough temporal expectancy in its own right, rendering the additional entrainment section redundant. Interestingly, the one task to show some positive effects of entrainment was NWR. Single words in isolation are thought too short to induce entrainment (Doelling, Arnal, Ghitza, & Poeppel, 2014; Ghitza, 2013) and so it could be that here the entrainment segment was able to ‘add value’ to the existing stimulus. For AMC children, the effect was strongest in unstressed syllables, with no effect on stressed syllables. This suggests the entrainment here successfully created a scaffolding structure for the more vulnerable unstressed syllables, facilitating better repetition. Significantly, in order for this to be the case, the internal structure of each word must have been represented (i.e. not just the stressed syllable rate). Entrainment at word level can therefore provide support for a word’s internal prosodic structure.

Regarding features intrinsic to the entrainment, we provided entrainment at both the stressed syllable and syllable level. If temporal expectancies operate maximally on only one level – i.e. at the stressed syllable rate, then the consistent stressed syllable beat will have done its job, and further information provided about the rhythm of the intervening syllables may have been superfluous. This may explain the lack of an accuracy effect for the Lexical Boundaries (LB) task – to make full use of the entraining beat, the rhythmic template at both stressed syllable and syllable levels would have to be abstracted and stored, whereas temporal expectancy would be created simply at the stressed syllable level. A similar phenomenon would potentially account for the lack of regularity effect (SR) – again stressed syllables were always at constant intervals regardless of the number of intervening weak syllables and so changes in the intervening weak syllable number may not have impacted performance due to the regular stressed beat.

The acoustic characteristics of entrainment may also have been relevant. The beat itself was digitally generated and so, although its general characteristics in terms of pitch, amplitude and duration were compatible with those of the stimulus, it was neither a human voice nor speech-like. It could be that using a more ecological entrainment segment, such as a human voice in single syllables (e.g. ‘ba’), would have provided a stronger result. Using a human voice would also have allowed us to signal larger prosodic units more readily (see below).

In the light of our findings regarding the importance of higher level prosodic phrasing patterns (e.g. at phrase level (SR) and clause and sentence level (LS, MS)), it could also be that our simplified entrainment lacked sufficient detail to cue these more complex aspects of prosody. We decided to use a simplified rhythm giving only syllable and stress level cues as we thought this would be the most direct. However, it could be that this simplified rhythm actually masked some of the cues relevant to higher level structures. In particular, we used a uniform acoustic configuration of amplitude, pitch and duration to indicate each strong and weak syllable. In contrast, in natural speech the degree of stress is indicated on a sliding, relative scale across longer units. By eliminating these relative cues, we may have inadvertently diminished the utility of the entrainment in anticipating the higher-level structures of the stimulus.

10.3 Acoustic Threshold (AT) Sensitivity

We had predicted that children with DLD would have larger thresholds for some of the acoustic cues to stress (Rise Time, Frequency, Duration) and that this would be related to task performance. In line with previous studies, we found that children with DLD had elevated thresholds for Rise Time (cf Corriveau, Pasquini, & Goswami, 2007; Richards & Goswami, 2015); Frequency (cf McArthur, Hogben, Edwards, Heath, & Mengler, 2000) and Duration (cf Corriveau et al., 2007; Cumming, Wilson, & Goswami, 2015; Richards & Goswami, 2015). We found that there was a significant correlational relationship between acoustic sensitivity and task performance for Rise Time (LB, NWR, LS, SR, MS), Frequency (LB, NWR, LS, SR, MS), Duration (NWR, LS, SR, MS) and Intensity (LB only). In general, these data support the idea that sensitive discrimination of acoustic cues to rhythm may relate to performance in language tasks. For some of the tasks, regressional analyses showed that age (LB, LS, SR) and NVIQ (NWR, SR, MS) contributed significant unique variance to individual differences in task performance. In some cases, this meant that the auditory predictors were no longer significant. Nevertheless, acoustic sensitivity accounted for additional significant variance in many of the task scores (NWR, LS, SR) once age and NVIQ had been accounted for.

The elevated thresholds we found for the DLD group may be of importance for understanding the ways in which the children responded to some of the rhythmic manipulations of the tasks. The AT tasks measure sensitivity to acoustic differences, and this is significant, since ‘stress’ is an emergent property of the surrounding acoustic landscape (Dilley & Pitt, 2010; Niebuhr, 2009) – stress levels are not absolutes, but exist only as contrasts within a specific context. Difficulties in detecting subtle differences between acoustic cues, regardless of absolute values, is therefore likely to lead to difficulties in determining the consistent relationships between stress levels required to abstract a relational template from a variety of absolute exemplars.

10.4 Children with DLD

Children with DLD scored more poorly in four of the five experimental tasks and had higher thresholds for discriminating Rise Time, Duration and Frequency than AMC children. The children with DLD therefore demonstrated difficulties in performing these rhythm-based language tasks.

However, the children with DLD were not impervious to the rhythmic manipulations employed in the tasks. They were able to differentiate between 2- and 3-item wordlists (LB), and made more accurate lexical stress judgements when the completing word was rhythmically congruent (LS). Children with DLD were also more accurate at repeating stressed syllables than unstressed (NWR and SR) and were sensitive to the metricality of stimuli (MS). Children with DLD are therefore capable of processing rhythm on some levels, but are less successful at doing so than their TD peers.

There are two complementary aspects of rhythm which the data implicate as problematic for children with DLD. One is difficulty in processing the relative nature of rhythm and stress, the other is in integrating individual units into wider rhythmic patterns, and with integrating these patterns with linguistic information.

Children with DLD were sensitive to some aspects of stress. At an individual syllable level, stressed syllables were more accurately repeated than unstressed (NWR and SR), indicating that stressed syllables occupy the same relatively privileged status in memory systems as for the AMC children. However, the results of tasks probing stress *patterning* revealed subtle profile differences between DLD and AMC children.

In SR, AMC children showed that the presence of stressed syllables in words supported whole-word repetition, possibly through providing a stable anchor to scaffold the whole-word structure. In contrast, this effect was not seen for the DLD children. This suggests that the DLD children may have been responding to the stressed syllable as an individual unit (hence higher stressed syllable accuracy rates) but not incorporating this with surrounding weak syllables into a larger word-level structure.

A lack of integration of units into larger patterns could also account for the lower score obtained in the LS task. This task required both processing of the stimulus pattern and comparison with a stored pattern. Poor online integration of the stimulus into a pattern, coupled with an underspecified stored pattern would make accurate judgement in this task problematic. This again suggests difficulty with patterning between stressed syllable and unstressed syllable levels.

Children with DLD also had difficulty in integrating rhythmic patterns with larger linguistic structures beyond the word. The MS task demonstrated that the children were able to respond to strong metrical structures, but did not relate these to the parallel syntactic structures. This seems to reflect a difficulty in integrating rhythmic and linguistic patterns. The metrical response indicates that children can perceive a strong, stable stressed syllable beat, however have not learnt how rhythmic and syntactic structures typically interact.

The difficulty in processing rhythmic *patterns* also brings us to the parallel issue of acoustic hierarchies. Perception of speech units is governed by the acoustic characteristics of each unit relative to the characteristics of surrounding units (Dilley & Pitt, 2010). Adult listeners retrospectively apply stress judgements to ambiguous utterance-initial syllables based on the stress characteristics of the subsequent utterance (Niebuhr, 2009) whilst judgements of utterance-final units are also affected by the characteristics of the preceding utterance (Dilley & McAuley, 2008). Since rhythm and stress are not absolutes, specifics will vary from speaker to speaker and utterance to utterance, so judgements cannot be made against a consistent benchmark, but must always be made on-line in a process of constant re-adjustment and re-evaluation.

Making rapid on-line judgements implies reliance on sensitive discrimination of the relevant acoustic cues. We know from the AT tasks that children with DLD have poorer sensitivity to Rise Time, Frequency and Duration than TD children, and so it seems likely that they will have greater difficulty in the constant evaluation of hierarchical acoustic patterns which may support efficient language processing.

A further implication of a hierarchical perspective is that the stored templates of rhythm and stress patterning must also be expressed in relative rather than absolute terms. Difficulty in discriminating relative characteristics is likely to lead to underspecified template representations. For individual units, poor sensitivity is likely to lead to difficulties in distinguishing between levels of stress unless the acoustic differentiation between levels is relatively marked. An Sw pattern, for example, may be erroneously processed as SS or ww due to acoustic characteristics not reaching a child's discrimination threshold. This further implies that children with poor discrimination may need an increased number of exposures to a given pattern for that pattern to become established. The more so since apparently undifferentiated exposures (e.g. ww above) would effectively be providing counter-examples. Poor discrimination of acoustic cues could therefore result in distorted perception and storage of rhythmic patterns.

We therefore see two interwoven areas of difficulty regarding language rhythm. Poor acoustic discrimination of cues to stress patterning is likely to lead to poorer abstraction of the relative stress patterns present in language. Difficulty in abstracting patterns is likely to be part of the difficulties seen in detecting word boundaries, building rhythmic frameworks for new vocabulary, rejecting incorrect lexical stress patterns and integrating rhythmic and sentence-level information.

As a practitioner, for me the question is not only of identifying what may be wrong, but also of identifying potential means for supporting language processing. Children with DLD showed some evidence of temporal expectancy (LS) and of following a stressed syllable beat (MS). In Large & Kolen (1994)'s terms this demonstrates that structured input can help children to determine *when?*. The entrainment effect we saw at Wordlength 3 (NWR) indicates that structured support can also help children to determine *what?* by scaffolding representations of stress patterns.

These insights could constitute a starting point for investigating possible interventions. In order to develop rhythmic patterning skills at word-level, for example, it would seem that using short words (no more than three syllables – NWR) is important, with stress differentiation sufficient to cross threshold. Presenting multiple exemplars whilst drawing explicit attention to the *patterning* could help secure stress pattern representations for target lexical items, and provide a scaffold for representing phonological detail. Presenting the target word within a strong, congruent stressed syllable beat should also facilitate this process (LS).

At a clausal level, children with DLD are sensitive to metrical structure (MS). This strength could be used to explicitly draw their attention to the relationship between rhythm and syntax. Examples from children's literature (such as used here) could be employed for this. This would enable children to draw on the *when?* of the strong beat to discover the *what?* of the syntactic structure.

10.5 Concluding Remarks

This study set out to investigate a hypothesised role for rhythmic sensitivity in children with DLD. We discovered that children with and without language disorder were affected by the rhythmic manipulations in our experiments, supporting the notion of a rhythmic influence in language processing.

Rhythmic structure at a lexical level affected NWR responses across groups, while congruency of CarrierPhrase and Target rhythm influenced judgements of lexical stress (AMC & DLD groups). Rhythmic structure at a lexical and phrasal level influenced responses in Sentence Repetition, whilst AMC children demonstrated awareness of the interface between rhythmic and syntactic structures (MS).

Stressed syllables were more accurately repeated than unstressed syllables for all groups (NWR & SR), with some evidence for the stressed syllable as the anchor point in phonological memory around which larger units are constructed (NWR & SR).

Provision of an entraining rhythm was not as supportive of task response as we had expected, although there was some evidence for facilitation of speed of response (LB – DLD group) and word-level scaffolding (NWR).

Children demonstrated a relationship between acoustic sensitivity to Rise Time (LB, NWR, LS, SR, MS), Frequency (LB, NWR, LS, SR, MS), Duration (NWR, LS, SR, MS), Intensity (LB) and scores on the respective experimental tasks, supporting a relationship between acoustic sensitivity and aspects of language rhythm. The DLD group had higher thresholds for Rise Time, Frequency and Duration (but did not differ for Intensity) and it is possible these higher thresholds could relate to the specific cross-task areas of difficulty found.

Children with DLD were able to respond to stressed syllables as individual units and showed some evidence of tracking the stressed syllable rate across time, but appeared to have difficulty in integrating these responses into larger rhythmic patterns, together with the corresponding linguistic units.

This profile of strengths and weaknesses in rhythmic sensitivity may account for some of the language difficulties these children experience and has the potential to be built upon in devising rhythmic interventions in the future.

Bibliography

Abecasis, D., Brochard, R., Granot, R., & Drake, C. (2005). Differential Brain Response to Metrical Accents in Isochronous Auditory Sequences. *Music Perception*, *22*(3), 549–562. http://doi.org/10.1525/mp.2005.22.3.549

Adams, A. M., & Gathercole, S. E. (1995). Phonological working memory and speech production in preschool children. *Journal of Speech, Language, and Hearing Research*, *38*(2), 403-414.

Adams, A. M., & Gathercole, S. E. (2000). Limitations in working memory: implications for language development. *International Journal of Language & Communication Disorders*, *35*(1), 95–116. http://doi.org/10.1080/136828200247278

Ahissar, E., Nagarajan, S., Ahissar, M., Protopapas, A., Mahncke, H., & Merzenich, M. M. (2001). Speech comprehension is correlated with temporal response patterns recorded from auditory cortex. *Proceedings of the National Academy of Sciences*, *98*(23), 13367–13372. http://doi.org/10.1073/pnas.201400998

Alcock, K. J., Passingham, R. E., Watkins, K., & Vargha-Khadem, F. (2000). Pitch and Timing Abilities in Inherited Speech and Language Impairment. *Brain and Language*, *75*, 34–46. http://doi.org/10.1006/brln.2000.2323

Alloway, T. P., & Gathercole, E. S. (2005). The role of sentence recall in reading and language skills of children with learning difficulties. *Learning and Individual Difference*, *15*(4), 271–282. http://doi.org/10.1017/CBO9781107415324.004

Alloway, T. P., Gathercole, S. E., Willis, C., & Adams, A. M. (2004). A structural analysis of working memory and related cognitive skills in young children. *Journal of Experimental Child Psychology*, *87*(2), 85-106.

Archibald, L. M. D., & Gathercole, S. E. (2006). Short-term and working memory in specific language impairment. *International Journal of Language and Communication Disorders. 41*(6), 675-693.

Archibald, L. M. D., & 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. http://doi.org/10.3758/BF03194122

Archibald, L. M. D., Gathercole, S. E., & Joanisse, M. F. (2009). Multisyllabic nonwords: more than a string of syllables. *The Journal of the Acoustical Society of America*, *125*(3), 1712–22. http://doi.org/10.1121/1.3076200

Archibald, L. M. D., & Joanisse, M. F. (2009). On the sensitivity and specificity of nonword repetition and sentence recall to language and memory impairments in children. *Journal of Speech, Language, and Hearing Research*, *52*(4), 899–914. http://doi.org/10.1044/1092-4388(2009/08-0099)

Arvaniti, A. (2009). Rhythm, timing and the timing of rhythm. *Phonetica*, *66*(1–2), 46–63. http://doi.org/10.1159/000208930

Attridge, D. (1982). *The Rhythms of English Poetry.* London: Pearson

Baddeley, A. (2003). Working memory and language: An overview. *Journal of Communication Disorders*, *36*(3), 189–208. http://doi.org/10.1016/S0021-9924(03)00019-4

Beattie, R. L., & Manis, F. R. (2013). Rise time perception in children with reading and combined reading and language difficulties. *Journal of Learning Disabilities*, *46*(3), 200-209. http://doi.org/10.1177/0022219412449421

Benasich, A. A., & Tallal, P. (1996). Auditory temporal processing thresholds, habituation, and recognition memory over the 1st year. *Infant Behavior and Development*, *19*(3), 339–357. http://doi.org/10.1016/S0163-6383(96)90033-8

Bishop, D. V. M. (1983). T*est for Reception of Grammar.* Manchester: Department of Psychology, University of Manchester.

Bishop, D. V., North, T., & Donlan, C. (1996). Nonword repetition as a behavioral marker for inherited language impairment: Evidence from a twin study. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *37*(4), 391–403. http://doi.org/10.1111/j.1469-7610.1996.tb01420.x

Bishop, D. V. M., Bright, P., James, C., Bishop, S. J., & Van der Lely, H. K. J. (2000). Grammatical SLI: A distinct subtype of developmental language impairment? *Applied Psycholinguistics*, *21*(2), 159–181. http://doi.org/10.1017/S0142716400002010

Botting, N., & Conti-Ramsden, G. (2001). Non-word repetition and language development in children with specific language impairment (SLI). *International Journal of Language and Communication Disorders*, *36*(4), 421–432. http://doi.org/10.1080/13682820110074971

Briscoe, T. (1989). Lexical access in connected speech recognition. *Proceedings of the 27th Annual meeting of the Association for Computational Linguistics*, 84–90. http://doi.org/10.3115/981623.981634

Brochard, R., Abecasis, D., Potter, D., Ragot, R., & Drake, C. (2003). The "Ticktock" of our Internal Clock: Direct Brain Evidence of Subjective Accents in Isochronous Sequences. *Psychological Science*, *14*(4), 362–366. http://doi.org/10.1111/1467-9280.24441

Brown, M., Salverda, A. P., Dilley, L. C., & Tanenhaus, M. K. (2011). Expectations from preceding prosody influence segmentation in online sentence processing. *Psychonomic Bulletin & Review*, *18*(6), 1189–1196. http://doi.org/10.3758/s13423-011-0167-9

Brown, M., Salverda, A. P., Dilley, L. C., & Tanenhaus, M. K. (2015). Metrical expectations from preceding prosody influence perception of lexical stress. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(2), 306.

Carlson, R., Elenius, K., Granström, B., & Hunnicutt, H. (1985). Phonetic and orthographic properties of the basic vocabulary of five European languages. *STL-QPSR 1*, *1985*, 63-94.

Chiat, S., & Hirson, A. (1987). From conceptual intention to utterance : A study of impaired language output in a child with developmental dysphasia. *British Journal of Disorders of Communication*, *22*(1), 37–64.

Chiat, S., & Roy, P. (2007). The Preschool Repetition Test : An Evaluation of Performance in. *Journal of Speech Language and Hearing Research*, *50*, 429–443.

Clarke, E. F. (1989). The perception of expressive timing in music. *Psychological Research*, *51*(1), 2-9.

Conti-Ramsden, G. (2003). Processing and Linguistic Markers in Young Children with Specific Language Impairment (SLI). *Journal of Speech, Language, and Hearing Research*, *46*, 1029–1037.

Conti-Ramsden, G., Botting, N., & Faragher, B. (2001). Psycholinguistic markers for specific language impairment. *Journal of Child Psychology and Psychiatry*, *42*(6), 741–748. http://doi.org/10.1111/1469-7610.00770

Conti-Ramsden, G., Knox, E., Botting, N., & Simkin, Z. (2002). Educational placements and National Curriculum Key Stage 2 test outcomes of children with a history of specific language impairment. *British Journal of Special Education*, *29*(2), 76–82. http://doi.org/10.1111/1467-8527.00244

Cooper, R. P., & Aslin, R. N. (1990). Preference for Infant-directed Speech in the First Month after Birth. *Child Development*, *61*(5), 1584–1595. http://doi.org/10.1111/j.1467-8624.1990.tb02885.x

Corriveau, K. H., & Goswami, U. (2009). Rhythmic motor entrainment in children with speech and language impairments : Tapping to the beat. *Cortex*, *45*(1), 119–130. http://doi.org/10.1016/j.cortex.2007.09.008

Corriveau, K., Pasquini, E., & Goswami, U. (2007). Basic auditory processing skills and specific language impairment: a new look at an old hypothesis. *J Speech Lang Hear Res*, *50*(3), 647–666. http://doi.org/10.1044/1092-4388(2007/046)

Cumming, R., Wilson, A., & Goswami, U. (2015). Basic auditory processing and sensitivity to prosodic structure in children with specific language impairments: a new look at a perceptual hypothesis. *Frontiers in Psychology*, *6*:972. http://doi.org/10.3389/fpsyg.2015.00972

Cumming, R., Wilson, A., Leong, V., Colling, L. J., & Goswami, U. (2015). Awareness of Rhythm Patterns in Speech and Music in Children with Specific Language Impairments. *Frontiers in Human Neuroscience*, *9*, 672. http://doi.org/10.3389/fnhum.2015.00672

Curtin, S., Mintz, T. H., & Christiansen, M. H. (2005). Stress changes the representational landscape: Evidence from word segmentation. *Cognition*, *96*(3), 233–262. http://doi.org/10.1016/j.cognition.2004.08.005

Cutler, A., (1996). Prosody and the Word Boundary Problem. In J. L. Morgan & K. Demuth (Eds), *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition* (pp87-101). Mahwah, NJ: Lawrence Erlbaum Associates.

Cutler, A., & Butterfield, S. (1992). Rhythmic cues to speech segmentation: Evidence from juncture misperception. *Journal of Memory and Language*, *31*(2), 218–236. http://doi.org/10.1016/0749-596X(92)90012-M

Cutler, A., & Carter, D. M. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language*, *2*, 133–142.

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1986). The syllable's differing role in the segmentation of French and English. *Journal of Memory and Language*, *25*, 385–400.

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1992). The monolingual nature of speech segmentation by bilinguals. *Cognitive Psychology*, *24*(3), 381–410. http://doi.org/10.1016/0010-0285(92)90012-Q

Cutler, A., & Otake, T. (1994). Mora or Phoneme? Further Evidence for Language-Specific Listening. *Journal of Memory and Language*, *33*(6), 824–844. http://doi.org/10.1006/jmla.1994.1039

Dauer, R. M. (1983). Stress-timing and syllable-timing reanalysed. *Journal of Phonetics, 11*(1), 51-62.

De Cara, B., & Goswami, U. (2003). Phonological neighbourhood density: Effects in a rhyme awareness task in five-year-old children. *Journal of Child Language*, *30*(3), 695-710.

DeCasper, A. J., & Fifer,W. P. (1980). Of human bonding: Newborns prefer their mothers' voices. *Science*, 208, 1174–1176.

Decasper, A. J., Lecanuet, J., & Busnel, Marie-Claire, Granier-Deferre, Carolyn, Maugeais, R. (1994). Fetal Reactions to Recurrent Maternal Speech. *Infant Behavior and Development*, *17*, 159–164.

DeCasper, A. J., & Spence, M. J. (1986). Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behavior and Development*, 9,133–150.

Devescovi, A., & Cristina Caselli, M. (2007). Sentence repetition as a measure of early grammatical development in Italian. *International Journal of Language & Communication Disorders*, *42*(2), 187–208. http://doi.org/10.1080/13682820601030686

Dilley, L. C., & McAuley, J. D. (2008). Distal prosodic context affects word segmentation and lexical processing. *Journal of Memory and Language*, *59*(3), 294–311. http://doi.org/10.1016/j.jml.2008.06.006

Dilley, L. C., & Pitt, M. A. (2010). Altering Context Speech Rate Can Cause Words to Appear or Disappear. *Psychological Science*, *21*(11), 1664–1670. http://doi.org/10.1177/0956797610384743

Dockrell, J., Lindsay, G., Palikara, O., & Cullen, M.A. (2007). *Raising the achievements of children and young people with specific speech and language difficulties and other special educational needs through school, to work and college. RR837.* Nottingham: DfES. Retrieved August 16, 2010, from: http://www.dcsf.gov.uk/research/data/uploadfiles/RR837.pdf

Dodd, B., Holm, A., Hua, Z., & Crosbie, S. (2003). Phonological development : a normative study of British English-speaking children. *Clinical Linguistics & Phonetics*, *17*(8), 617–643. http://doi.org/10.1080/0269920031000111348

Doelling, K. B., Arnal, L. H., Ghitza, O., & Poeppel, D. (2014). Acoustic landmarks drive delta-theta oscillations to enable speech comprehension by facilitating perceptual parsing. *NeuroImage*, *85*, 761–768. http://doi.org/10.1016/j.neuroimage.2013.06.035

Doelling, K. B., & Poeppel, D. (2015). Cortical entrainment to music and its modulation by expertise. *Proceedings of the National Academy of Sciences*, *112*(45), E6233-E6242. http://doi.org/10.1073/pnas.1508431112

Dollaghan, C., & Campbell, T. (1998). Nonword repetition and child language impairment. *Journal of Speech, Language, and Hearing Research*, *41*(5), 1136–1146.

Donaldson, J. (2002). *Room on the Broom.* London: Pan Macmillan.

Dresher, B. Elan (1996). Introduction to Metrical and Prosodic Phonology. In J. L. Morgan & K. Demuth (Eds), *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition* (pp41-54). Mahwah, NJ: Lawrence Erlbaum Associates.

Drullman, R., Festen, J. M., & Plomp, R. (1994). Effect of reducing slow temporal modulations on speech reception. *The Journal of the Acoustical Society of America*, *95*(5), 2670–2680. http://doi.org/10.1121/1.409836

Dunn, L., Dunn, L., Whetton, C., & Burley, J. (1997). *British Picture Vocabulary Scales (*2nd ed.). London: NFER-Nelson.

Edwards, J., & Lahey, M. (1998). Nonword repetitions of children with specific language impairment: Exploration of some explanations for their inaccuracies. *Applied Psycholinguistics*, *19*(2), 279-309. http://doi.org/10.1017/S0142716400010079

Eimas, P. (1996). The Perception and Representation of Speech by Infants. In J. L. Morgan & K. Demuth (Eds), *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition* (pp25-40). Mahwah, NJ: Lawrence Erlbaum Associates.

Elliott, C. D., Smith, P., & McCulloch, K. (1996). *British Ability Scales – 2nd Edition*. Windsor: NFER-Nelson.

Ellis Weismer, S., Tomblin, J. B., Zhang, X., Buckwalter, P., Chynoweth, J. G., Jones, M. (2000). Nonword repetition perfomance in school-age children with and without language impairment. *Journal of Speech, Language, and Hearing Research, 43*(4), 865-878

Everitt, A., Hannaford, P., & Conti-Ramsden, G. (2013). Markers for persistent specific expressive language delay in 3–4-year-olds. *International Journal of Language & Communication Disorders*, *48*(5), 534-553.

Fraser, J., Goswami, U., & Conti-Ramsden, G. (2010). Dyslexia and Specific Language Impairment : The Role of Phonology and Auditory Processing. *Scientific Studies of Reading*, *14*(1), 8–29.

Frazier, L., Carlson, K., & Clifton, C. (2006). Prosodic phrasing is central to language comprehension. *Trends in Cognitive Sciences*, *10*(6), 244–249. http://doi.org/10.1016/j.tics.2006.04.002

Frederickson, N., Frith, U., & Reason, R. (1997). *Phonological Assessment Battery*. London: GL Assessment Limited.

Fry, D. B. (1958). Experiments in the perception of stress. *Language and Speech*, *1*(2), 126–152. http://doi.org/10.1177/002383095800100207

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

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

Gathercole, S. E., & Baddeley, A. D. (1989). Evaluation of the Role of Phonological STM in the Development Vocabulary in Children : A Longitudinal Study. *Journal of Memory and Language*, *28*, 200–213.

Gathercole, S. E., & Baddeley, A. D. (1990). Phonological memory deficits in language disordered children: Is there a causal connection? *Journal of Memory and Language*, *29*(3), 336–360. http://doi.org/10.1016/0749-596X(90)90004-J

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

Gathercole, S. E., Willis, C., Emslie, H., & Baddeley, A. D. (1991). The influences of number of syllables and wordlikeness on children's repetition of nonwords. Applied Psycholinguistics, *12*(3), 349. http://doi.org/10.1017/S0142716400009267

Gerken, L. (1991). The Metrical Basis for Children's Subjectless Sentences. *Journal of Memory and Language*, *30*, 431–451.

Gerken, L. (1994a). A metrical template account of children's weak syllable omissions from multisyllabic words. *Journal of Child Language*, *21*(3), 565–584. http://doi.org/10.1017/S0305000900009466

Gerken, L. (1994b). Young Children's Representation of Prosodic Phonology: Evidence from English-Speakers' Weak Syllable Productions. *Journal of Memory and Language*, *33*, 19–38.

Gerken, L. (1996). Prosody's Role in Language Acquisition and Adult Parsing. *Journal of Psycholinguistic Research*, *25*(2), 345–356. http://doi.org/10.1007/BF01708577

Gerken, L. A., Landau, B., & Remez, R. E. (1990). Function morphemes in young children's speech perception and production. *Developmental Psychology*, *26*(2), 204–216.

Ghitza, O. (2011). Linking speech perception and neurophysiology: Speech decoding guided by cascaded oscillators locked to the input rhythm. *Frontiers in Psychology*, *2*, 1–13. http://doi.org/10.3389/fpsyg.2011.00130

Ghitza, O. (2012). On the role of theta-driven syllabic parsing in decoding speech: Intelligibility of speech with a manipulated modulation spectrum. *Frontiers in Psychology*, *3*, 1–12. http://doi.org/10.3389/fpsyg.2012.00238

Ghitza, O. (2013). The theta-syllable: A unit of speech information defined by cortical function. *Frontiers in Psychology*, *4*(MAR), 1–5. http://doi.org/10.3389/fpsyg.2013.00138

Ghitza, O., Giraud, A., & Poeppel, D. (2013). Neuronal oscillations and speech perception: critical-band temporal envelopes are the essence. *Frontiers in Human Neuroscience*, *6*, 340. http://doi.org/10.3389/fnhum.2012.00340

Ghitza, O., & Greenberg, S. (2009). On the Possible Role of Brain Rhythms in Speech Perception : Intelligibility of Time-Compressed Speech with Periodic and Aperiodic Insertions of Silence. *Phonetica*, *66*, 113–126. http://doi.org/10.1159/000208934

Gordon, R. L., Shivers, C. M., Wieland, E. A., Kotz, S. A., Yoder, P. J., & Devin Mcauley, J. (2014). Musical rhythm discrimination explains individual differences in grammar skills in children. *Developmental Science*, *18*(4), 635–644. http://doi.org/10.1111/desc.12230

Goswami, U. (2011). A temporal sampling framework for developmental dyslexia. *Trends in Cognitive Sciences*, *15*(1), 3-10. http://doi.org/10.1016/j.tics.2010.10.001

Goswami, U., Gerson, D., & Astruc, L. (2010). Amplitude envelope perception, phonology and prosodic sensitivity in children with developmental dyslexia. *Reading and Writing*, *23*(8), 995–1019. http://doi.org/10.1007/s11145-009-9186-6

Goswami, U., Thomson, J., Richardson, U., Stainthorp, R., Hughes, D., Rosen, S., & Scott, S. K. (2002). Amplitude envelope onsets and developmental dyslexia: A new hypothesis. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(16), 10911–10916. http://doi.org/10.1073/pnas.122368599

Grabe, E., & Low, E. L. (2002). Durational variability in speech and the rhythm class hypothesis. In C. Gussenhoven & N. Warner (eds.). *Laboratory phonology* (Vol. 7, pp. 515–546). Berlin: Mouton de Gruyter.

Greenberg, S., Carvey, H., Hitchcock, L., & Chang, S. (2002). Beyond the phoneme: A juncture-accent model of spoken language. In *Proceedings of the second international conference on human language technology research* (pp. 36-43). Morgan Kaufmann Publishers Inc..

Greenberg, S., Carvey, H., Hitchcock, L., & Chang, S. (2003). Temporal properties of spontaneous speech - A syllable-centric perspective. *Journal of Phonetics*, *31*(3–4), 465–485. http://doi.org/10.1016/j.wocn.2003.09.005

Gross, J., Hoogenboom, N., Thut, G., Schyns, P., Panzeri, S., Belin, P., & Garrod, S. (2013). Speech Rhythms and Multiplexed Oscillatory Sensory Coding in the Human Brain. *PLoS Biology*, *11*(12): e1001752. http://doi.org/10.1371/journal.pbio.1001752

Harrington, J., Watson, G., & Cooper, M. (1989). Word boundary detection in broad class and phoneme strings. *Computer Speech and Language*, *3*(4), 367–382. http://doi.org/10.1016/0885-2308(89)90004-1

Hirsh-Pasek, K., Kemler Nelson, D. G., Jusczyk, P. W., Cassidy, K. W., Druss, B., & Kennedy, L. (1987). Clauses are perceptual units for young infants. *Cognition*, *26*(3), 269–286. http://doi.org/10.1016/S0010-0277(87)80002-1

Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review*, *96*(3), 459–491. http://doi.org/10.1037/0033-295X.96.3.459

Jones, M. R., Moynihan, H., MacKenzie, N., & Puente, J. (2002). Temporal Aspects of Stimulus-Driven Attending in Dynamic Arrays. *Psychological Science*, *13*(4), 313–319. http://doi.org/10.1111/1467-9280.00458

Jusczyk, P. W., Cutler, A., & Redanz, N. J. (1993). Infants' Preference for the Predominant Stress Patterns of English Words. *Child Development*, *64*, 675–687.

Jusczyk, P. W., Hirsh-Pasek, K., Kemler Nelson, D. G., Kennedy, L. J., Woodward, A., & Piwoz, J. (1992). Perception of acoustic correlates of major phrasal units by young infants. *Cognitive Psychology*, *24*(2), 252–293. http://doi.org/10.1016/0010-0285(92)90009-Q

Jusczyk, P. W., Houston, D. M., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology*, *39*(3–4), 159–207. http://doi.org/10.1006/cogp.1999.0716

Kan, P. F., & Windsor, J. (2010). Word Learning in Children With Primary Language Impairment: A Meta-Analysis. *Journal of Speech Language and Hearing Research*, *53*(3), 739-756. http://doi.org/10.1044/1092-4388(2009/08-0248)

Kemler Nelson, D. G., Hirsh-Pasek, K., Jusczyk, P. W., & Cassidy, K. W. (1989). How the prosodic cues in motherese might assist language learning. *Journal of Child Language*, *16*, 55–68.

Kessler, B., & Treiman, R. (1997). Syllable Structure and the Distribution of Phonemes in English Syllables. *Journal of Memory and Language*, *311*(37), 295–311.

Klem, M., Melby-Lervåg, M., Hagtvet, B., Lyster, S. A. H., Gustafsson, J. E., & Hulme, C. (2015). Sentence repetition is a measure of children's language skills rather than working memory limitations. *Developmental Science*, *18*(1), 146–154. http://doi.org/10.1111/desc.12202

Kochanski, G., Grabe, E., Coleman, J., & Rosner, B. (2005). Loudness predicts prominence: Fundamental frequency adds little. *Journal of the Acoustical Society of America*, 118, 1038-1054.

Kotz, S. A., Gunter, T. C., & Wonneberger, S. (2005). The basal ganglia are receptive to rhythmic compensation during auditory syntactic processing: ERP patient data. *Brain and Language*, *95*(1), 70–71. http://doi.org/10.1016/j.bandl.2005.07.039

Kotz, S. A., Schwartze, M., & Schmidt-Kassow, M. (2009). Non-motor basal ganglia functions: A review and proposal for a model of sensory predictability in auditory language perception. *Cortex*, *45*(8), 982–990. http://doi.org/10.1016/j.cortex.2009.02.010

Lakatos, P., Ankoor, S. S., Knuth, K. H., Ulbert, I., Karmos, G., & Schroeder, E. (2005). An Oscillatory Hierarchy Controlling Neuronal Excitability and Stimulus Processing in the Auditory Cortex. *Journal of Neurophysiology*, *94*, 1904–1911. http://doi.org/10.1152/jn.00263.2005

Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., & Schroeder, C. E. (2008). Entrainment of Neuronal Oscillations as a Mechanism of Attentional Selection. *Science*, *320*(5872), 110–113. http://doi.org/10.1126/science.1154735

Large, E. W., & Jones, M. R. (1999). The Dynamics of Attending: How People Track Time-Varying Events. *Pscyhological Review*, *106*(1), 119–159.

Large, E. W., & Kolen, J. F. (1994). Resonance and the perception of musical meter. *Connection Science*, *6*(2-3), 177–208. http://doi.org/10.1080/09540099408915723

Laver, J. (1994). *Principles of Phonetics*. Cambridge: Cambridge University Press.

Law, J., Rush, R., Schoon, I., & Parsons, S. (2009). Modeling Developmental Language Difficulties From School Entry Into Adulthood: Literacy, Mental Health, and Employment Outcomes. *Journal of Speech, Language, and Hearing Research*, *52,* 1401-1416. doi:10.1044/1092-4388(2009/08-0142)

Leech, G., Rayson, P., & Wilson, A. (2001). *Word Frequencies in Written and Spoken English*. Harlow: Pearson Education Ltd.

Leonard, L. (2000). Specific Language Impairment across languages. In D. Bishop & L. Leonard (Eds), *Speech and Language Impairments in Children* (pp115-131). Hove: Psychology Press.

Leonard, L. B., Ellis Weismer, S., Miller, C. A, Francis, D. J., Tomblin, J. B., & Kail, R. V. (2007). Speed of processing, working memory, and language impairment in children. *Journal of Speech, Language & Hearing Research*, *50*(2), 408–428. http://doi.org/10.1044/1092-4388(2007/029)

Leong, V. (2012). Prosodic rhythm in the speech amplitude envelope: Amplitude modulation phase hierarchies (AMPHs) and AMPH models. PhD Thesis, University of Cambridge, UK.

Leong, V., & Goswami, U. (2014). Impaired extraction of speech rhythm from temporal modulation patterns in speech in developmental dyslexia. *Frontiers in Human Neuroscience*, *8*, 1–14. http://doi.org/10.3389/fnhum.2014.00096

Leong, V., Kalashnikova, M., Burnham, D., & Goswami, U. (2014). Infant-directed speech enhances temporal rhythmic structure in the envelope. *INTERSPEECH 2014, 15th Annual Conference of the International Speech Communication Association (ISCA).*

Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge, MA: MIT Press.

Lombardi, L., & Potter, M. C. (1992). The Regeneration of Syntax in Short Term Memory. *Journal of Memory and Language*, *31*, 713–733.

Luo, H., & Poeppel, D. (2007). Phase Patterns of Neuronal Responses Reliably Discriminate Speech in Human Auditory Cortex. *Neuron*, *54*(6), 1001–1010. http://doi.org/10.1016/j.neuron.2007.06.004

Luo, H., & Poeppel, D. (2012). Cortical oscillations in auditory perception and speech: Evidence for two temporal windows in human auditory cortex. *Frontiers in Psychology*, *3*(MAY), 1–10. http://doi.org/10.3389/fpsyg.2012.00170

Marchman, V. A., Wulfeck, B., & Ellis Weismer, S. (1999). Morphological productivity in children with normal language and SLI: A study of the English past tense. *Journal of Speech, Language, and Hearing Research,* 42 (1), 206-219.

Marshall, C. M., & Nation, K. (2003). Individual differences in semantic and structural errors in children's memory for sentences. *Educational and Child Psychology*, *20*(3), 7-18.

Martin, J. G. (1970). Rhythm-Induced Judgments of Word Stress in Sentences. *Journal of Verbal Learning and Verbal Behavior*, *9*, 627–633.

Marton, K., & Schwartz, R. G. (2003). Working memory capacity and language processes in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *46*(5), 1138–1153. http://doi.org/10.1044/1092-4388(2003/089)

McArthur, G. M., & Bishop, D. V. M. (2004). Which People with Specific Language Impairment have Auditory Processing Deficits? *Cognitive Neuropsychology*, *21*(1), 79–94. http://doi.org/10.1080/02643290342000087

McArthur, G. M., & Hogben, J. H. (2001). Auditory backward recognition masking in children with a specific language impairment and children with a specific reading disability. *The Journal of the Acoustical Society of America*, *109*(3), 1092–1100. http://doi.org/10.1121/1.1338559

McArthur, G. M., Hogben, J. H., Edwards, V. T., Heath, S. M., & Mengler, E. D. (2000). On the "Specifics" of Specific Reading Disability and Specific Language Impairment. *Journal of Child Psychology and Psychiatry*, *41*(7), 869–874. http://doi.org/10.1111/1469-7610.00674

McGregor, K. K., & Leonard, L. B. (1994). Subject pronoun and article omissions in the speech of children with specific language impairment: A phonological interpretation. *Journal of Speech, Language, and Hearing Research*, *37*(1), 171-181.

Mehler, J., Dupoux, E., Nazzi, T., & Deheaene-Lambertz, G. (1996). Coping with Linguistic Diversity: The Infant's Viewpoint. In J. L. Morgan & K. Demuth (Eds), *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition* (pp101-117). Mahwah, NJ: Lawrence Erlbaum Associates.

Mehler, J., Jusczyk, P., Lambertz, G., & Halsted, N. (1988). A precursor of language acquisition in young infants. *Cognition*, *29*, 143–178.

Mengler, E. D., Hogben, J. H., Michie, P., & Bishop, D. V. M. (2005). Poor Frequency Discrimination is related to Oral Language Disorder in Children: A Psychoacoustic Study. *Dyslexia*, *11*, 155–173. http://doi.org/10.1002/dys.302

Mines, A. M., Hanson, B. F., & Shoup, J. E. (1978). Frequency of Occurrence of Phonemes in Conversational English. *Language and Speech*, *21*(3), 221–241.

Morgan, J., & Demuth, K. (1996). Signal to Syntax: An Overview. In J. L. Morgan & K. Demuth (Eds), *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition* (pp1-25). Mahwah, NJ: Lawrence Erlbaum Associates.

Morgan, J. L., & Saffran, J. R. (1995). Emerging Integration of Sequential and Suprasegmental Information in Preverbal Speech Segmentation. *Child Development*, *66*(4), 911–936.

MRC Psycholinguistic Database. *Machine Usable Dictionary, Version 2.00*. Retrieved from http://websites.psychology.uwa.edu.au/school/MRCDatabase/uwa_mrc.htm

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*, 1033–1047. http://doi.org/10.1044/1092-4388(2005/072)

Murdock, B. B., (1962). The serial position effect of free recall. *Journal of Experimental Psychology ,* *64*(5), 482-488.

Nash, H. M., Hulme, C., Gooch, D., & Snowling, M. J. (2013). Preschool language profiles of children at family risk of dyslexia: Continuities with specific language impairment. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *54*(9), 958–968. http://doi.org/10.1111/jcpp.12091

Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: toward an understanding of the role of rhythm. *Journal of Experimental Psychology. Human Perception and Performance*, *24*(3), 756–766. http://doi.org/10.1037/0096-1523.24.3.756

Nespor, M., & Vogel, I. (1986). Prosodic phonology. *Dordrecht: Foris*.

Newman, R., Ratner, N. B., Jusczyk, A. M., Jusczyk, P. W., & Dow, K. A. (2006). Infants' early ability to segment the conversational speech signal predicts later language development: a retrospective analysis. *Developmental Psychology*, *42*(4), 643–655. http://doi.org/10.1037/0012-1649.42.4.643

Niebuhr, O. (2009). F0-based rhythm effects on the perception of local syllable prominence. *Phonetica*, *66*(1–2), 95–112. http://doi.org/10.1159/000208933

Norbury, C. F., Gooch, D., Wray, C., Baird, G., Charmand, T., Simonoff, E., Vamvakas, G., & Andrew, P. (2016). The impact of NVIQ on prevalence and clinical presentation of language disorder: evidence from a population study. *Journal of Child Psychology and Psychiatry*, 1–37. http://doi.org/10.1111/jcpp.12573

OED. (2001). metre | meter, n. Retrieved from Oxford University Press. http://www.oed.com/view/Entry/117646?rskey=y05Mfo&result=1&isAdvanced=false#eid

Panzeri, S., Brunel, N., Logothetis, N. K., & Kayser, C. (2010). Sensory neural codes using multiplexed temporal scales. *Trends in Neurosciences*, *33*(3), 111–120. http://doi.org/10.1016/j.tins.2009.12.001

Patel, A. (2008). *Music, Language and the Brain.* Oxford: OUP

Peelle, J. E., & Davis, M. H. (2012). Neural oscillations carry speech rhythm through to comprehension. *Frontiers in Psychology*, *3*, 1–17. http://doi.org/10.3389/fpsyg.2012.00320

Peppé, S., & McCann, J. (2003). *Profiling Elements of Prosody in Speech-Communication (PEPS-C*). Retrieved from http://www.peps-c.com/index.html

Poeppel, D. (2003). The analysis of speech in different temporal integration windows: Cerebral lateralization as “asymmetric sampling in time.” *Speech Communication*, *41*(1), 245–255. http://doi.org/10.1016/S0167-6393(02)00107-3

Poeppel, D., Idsardi, W. J., & van Wassenhove, V. (2008). Speech perception at the interface of neurobiology and linguistics. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *363*(1493), 1071–1086. http://doi.org/10.1098/rstb.2007.2160

Polišenská, K., Chiat, S., & Roy, P. (2015). Sentence repetition: What does the task measure? *International Journal of Language and Communication Disorders*, *50*(1), 106–118. http://doi.org/10.1111/1460-6984.12126

Polka, L., Sundara, M., & Blue, S. (2002). The role of language experience in word segmentation: A comparison of English, French, and bilingual infants. *The Journal of the Acoustical Society of America*, *111*(5), 2455-2455.

Potter, C. M., & Lombardi, L. (1990). Regeneration in the short term recall of sentences. *Journal of Memory and Language*, *29*, 633–654.

Power, A. J., Mead, N., Barnes, L., & Goswami, U. (2012). Neural entrainment to rhythmically presented auditory, visual, and audio-visual speech in children. *Frontiers in Psychology*, *3*(2), 106–109. http://doi.org/10.3389/fpsyg.2012.00216

Price, P. J., Ostendorf, M., Shattuck-Hufnagel, S., & Fong, C. (1991). The use of prosody in syntactic disambiguation. *The Journal of the Acoustical Society of America*, *90*(6), 2956–2970. http://doi.org/10.1121/1.401770

Przybylski, L., Bedoin, N., Krifi-Papoz, S., Herbillon, V., Roch, D., Léculier, L., Kotz, S. A., & Tillmann, B. (2013). Rhythmic auditory stimulation influences syntactic processing in children with developmental language disorders. *Neuropsychology*, *27*(1), 121–131. http://doi.org/10.1037/a0031277

Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition*, *73*(3), 265–292. http://doi.org/10.1016/S0010-0277(99)00058-X

Repp, B. H. (1992). Probing the cognitive representation of musical time: Structural constraints on the perception of timing perturbations. *Cognition*, *44*, 241–281.

Richards, S. (2010). *Stress, Rhythm and Pitch in Language and Music Processing in Specific Language Impairment*. M.Ed. thesis, University of Cambridge, UK.

Richards, S., & Goswami, U. (2015). Auditory Processing in Specific Language Impairment (SLI): Relations With the Perception of Lexical and Phrasal Stress. *Journal of Speech and Hearing Research*, *58*, 1292–1305.

Riches, N. G., Loucas, T., Baird, G., Charman, T., & Simonoff, E. (2010). Sentence repetition in adolescents with specific language impairments and autism: An investigation of complex syntax. *International Journal of Language & Communication Disorders*, *45*(1), 47-60.

Riches, N. G. (2012). Sentence repetition in children with specific language impairment: An investigation of underlying mechanisms. *International Journal of Language and Communication Disorders*, *47*(5), 499–510. http://doi.org/10.1111/j.1460-6984.2012.00158.x

Rohl, M., & Pratt, C. (1995). Phonological awareness, verbal working memory and the acquisition of literacy. *Reading and Writing*, *7*(4), 327–360. http://doi.org/10.1007/BF01027723

Rosen, S. (1992). Temporal information in speech: acoustic, auditory and linguistic aspects. *Philosophical Transactions of the Royal Society of London B: Biological Sciences,* 336 , 367-373 http://doi.org/10.1098/rstb.1992.0070

Roy, P., & Chiat, S. (2004). A prosodically controlled word and nonword repetition task for 2- to 4-year-olds: evidence from typically developing children. *Journal of Speech, Language, and Hearing Research*, *47*(1), 223–234. http://doi.org/10.1044/1092-4388(2004/019)

Sansavini, A., Bertoncini, J., & Giovanelli, G. (1997). Newborns discriminate the rhythm of multisyllabic stressed words. *Developmental Psychology*, *33*(1), 3–11. http://doi.org/10.1037//0012-1649.33.1.3

Schmidt-Kassow, M., & Kotz, S. A. (2008). Entrainment of syntactic processing? ERP-responses to predictable time intervals during syntactic reanalysis. *Brain Research*, *1226*, 144-155.

Schmidt-Kassow, M., & Kotz, S. A. (2009). Event-related brain potentials suggest a late interaction of meter and syntax in the P600. *Journal of Cognitive Neuroscience*, *21*(9), 1693-1708.

Scholes, R. J. (1970). On functors and contentives in children's imitations of word strings. *Journal of Verbal Learning and Verbal Behavior*, *9*(2), 167–170. http://doi.org/10.1016/S0022-5371(70)80046-9

Schroeder, C. E., & Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends in Neurosciences*, *32*(1), 9–18. http://doi.org/10.1016/j.tins.2008.09.012

Seeff-Gabriel, B., Chiat, S., & Roy, P. (2008). *Early Repetition Battery*. London: Pearson

Seeff-Gabriel, B., Chiat, S., & Dodd, B. (2010). Sentence imitation as a tool in identifying expressive morphosyntactic difficulties in children with severe speech difficulties. *International Journal of Language & Communication Disorders*, *45*(6), 691–702. http://doi.org/10.3109/13682820903509432

Segalowitz, S. J., & Lane, K. C. (2000). Lexical access of function versus content words. *Brain and Language, 75*, 376-389.

Selkirk, E. (1996). The Prosodic Structure of Function Words. In J. L. Morgan & K. Demuth (Eds), *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition* (pp187-215). Mahwah, NJ: Lawrence Erlbaum Associates.

Semel, E., Wiig, E. H., & Secord, W. A. (2000). *Clinical Evaluation of Language Fundamentals – third edition*, *UK Version (CELF-3^{UK}).* San Antonio, TX: The Psychological Corporation.

Shannon, R. V, Zeng, F-G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science*, *270*(5234), 303–304. http://doi.org/10.1126/science.270.5234.303

Smith, Z. M., Delgutte, B., & Oxenham, A. J. (2002). Chimaeric sounds reveal dichotomies in auditory perception. *Nature*, *416*(6876), 87–90.

Smolík, F., & Vávrů, P. (2014). Sentence imitation as a marker of SLI in Czech: Disproportionate impairment of verbs and clitics. *Journal of Speech, Language, and Hearing Research*, *57*(3), 837-849.

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

Snowling, M. J., Duff, F. J., Nash, H. M., & Hulme, C. (2015). Language profiles and literacy outcomes of children with resolving, emerging, or persisting language impairments. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 57*(12), 1360-1369. http://doi.org/10.1111/jcpp.12497

Stokes, S. F., Wong, A. M-Y., Fletcher, P., & Leonard, L. B. (2006). Nonword Repetition and Sentence Repetition as Clinical Markers of Specific Language Impairment: The Case of Cantonese. *Journal of Speech Language and Hearing Research*, *49*(2), 219-236. http://doi.org/10.1044/1092-4388(2006/019)

Tallal, P., & Piercy, M. (1973). Defects of non-verbal auditory perception in children with developmental aphasia. *Nature*, *241*(5390), 468–469. http://doi.org/10.1038/241468a0

Thomassen, J. (1982). Melodic accent: Experiments and a tentative model. *Journal of the Acoustical Society of America*, *71*(6), 1598–1605. http://dx.doi.org/10.1121/1.387814

Thut, G., Miniussi, C., & Gross, J. (2012). The functional importance of rhythmic activity in the brain. *Current Biology*, *22*(16), R658–R663. http://doi.org/10.1016/j.cub.2012.06.061

Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech, Language, and Hearing Research*, *40*(6), 1245–1260.

Tomblin, J., B., Zhang, X., Buckwalter, P., & Catts, H. (2000). The association of reading disability, behavioral disorders and language impairment among second-grade children, *Journal of Child Psychology and Psychiatry and Allied Disciplines, 41*(4), 473-482

van der Lely, H., & Christian, V. (2000). Lexical word formation in grammatical SLI children: A grammar-specific versus an input-processing deficit? *Cognition*, *75*, 33–63.

van der Lely, H. K. J., Rosen, S., & Adlard, A. (2004). Grammatical language impairment and the specificity of cognitive domains: Relations between auditory and language abilities. *Cognition*, *94*(2), 167–183. http://doi.org/10.1016/j.cognition.2004.01.003

van der Lely, H., Rosen, S., & McClelland, A. (1998). Evidence for a grammar-specific deficit in children. *Current Biology*, *8*(23), 1253–1258. http://doi.org/10.1016/S0960-9822(07)00534-9

van Ooijen, B., Bertoncini, J., Sansavini, a, & Mehler, J. (1997). Do weak syllables count for newborns? *The Journal of the Acoustical Society of America*, *102*(6), 3735–3741. http://doi.org/10.1121/1.420136

von Bülow, H. cited in Watson, D. (1994). *The Wordsworth dictionary of musical quotations*. Ware: Wordsworth Editions.

Weber, C., Hahne, A., Friedrich, M., & Friederici, A. D. (2005). Reduced stress pattern discrimination in 5-month-olds as a marker of risk for later language impairment: Neurophysiologial evidence. *Cognitive Brain Research*, *25*(1), 180–187. http://doi.org/10.1016/j.cogbrainres.2005.05.007

Wechsler, D. (1992). *Wechsler intelligence scale for children-third edition (WISC-III)*. San Antonio, Texas: The Psychological Corporation.

Weinert, S. (1992). Deficits in acquiring language structure: the importance of using prosodic cues. *Applied Cognitive Psychology*, 6, 545-571

Wells, B., & Peppé, S. (2003). Intonation Abilities of Children With Speech and Language Impairments. *Journal of Speech, Language, and Hearing Research*, *46*(1), 5–20. http://doi.org/10.1044/1092-4388(2003/001)

Wells, B., Peppé, S., & Goulandris, N. (2004). Intonation development from five to thirteen. *Journal of Child Language*, *31*(4), 749–778. http://doi.org/10.1017/S030500090400652X

Wenk, B. J., & Wioland, F. (1982). Is French really syllable-timed? *Journal of Phonetics, 10(*2)*,* 193-216.

Willis, C. S., & Gathercole, S. E. (2001). Phonological short-term memory contributions to sentence processing in young children. *Memory*, *9*(4), 349–363. http://doi.org/10.1080/09658210143000155

Appendix A – Experimental Tasks

Table A-1 List of experimental tasks completed by each participant group.

Task	**AMC**	**DLD**	**YLC**
Acoustic Threshold Estimation			
Frequency	√	√	-
Duration	√	√	-
Rise Time	√	√	-
Intensity	√	√	-
Experimental			
Lexical Boundaries – Unentrained	√	√	√
Lexical Boundaries – Entrained	√	√	-
Lexical Stress – Unentrained	√	√	√
Lexical Stress - Entrained	√	√	-
Nonword Repetition - Unentrained	√	√	√
Nonword Repetition - Entrained	√	√	-
Sentence Repetition – Unentrained	√	√	√
Sentence Repetition - Entrained	√	√	-
Metrical Stress 1	√	√	√
Metrical Stress 2	√	√	√
Metrical Stress 3	√	√	√

Appendix B - Scatterplots for Acoustic Thresholds and Task Scores

1. Experimental Task 1 - Lexical Boundaries

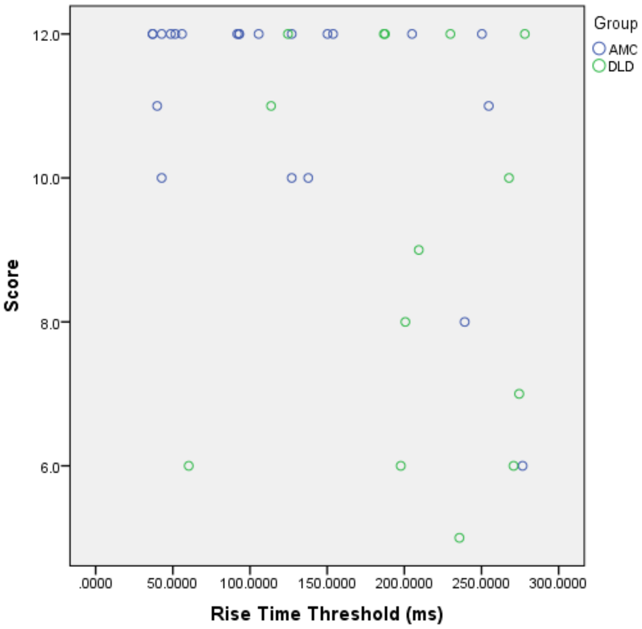


Figure B-1-1 Scatterplot of Rise Time threshold by Score (AMC and DLD groups)

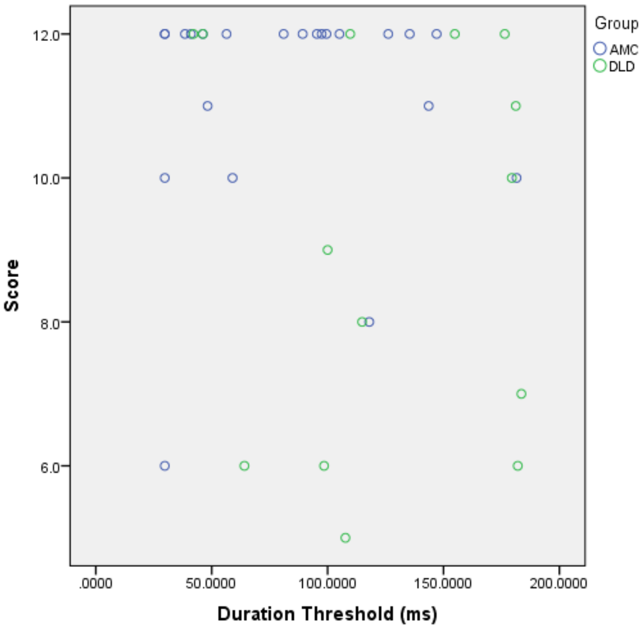


Figure B-1-2 Scatterplot of Duration threshold by Score (AMC and DLD groups)

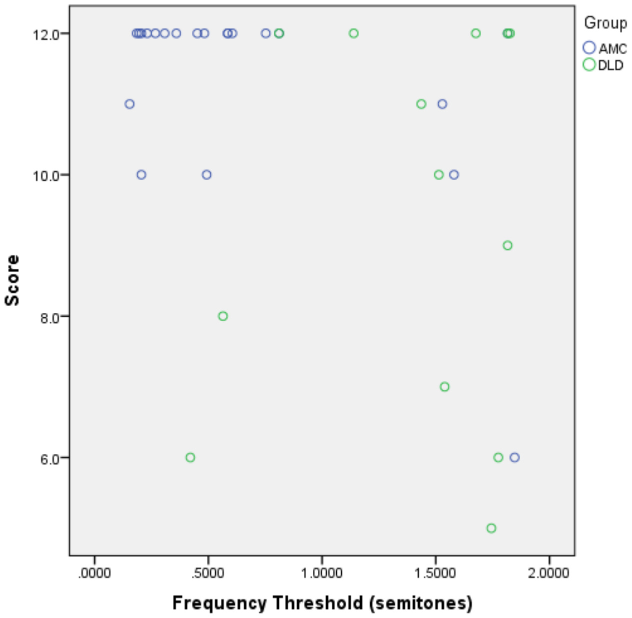


Figure B-1-3 Scatterplot of Frequency threshold by Score (AMC and DLD groups)

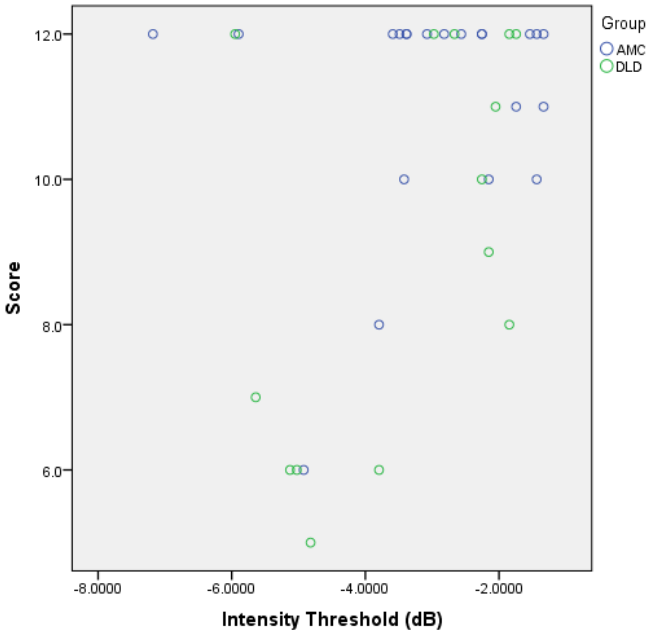


Figure B-1-4 Scatterplot of Intensity threshold by Score (AMC and DLD groups)

2. Experimental Task 2 – Nonword Repetition

i) Word Score

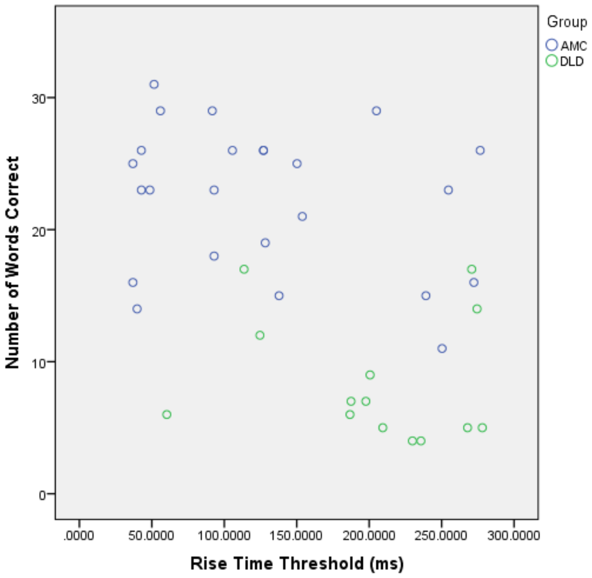


Figure B-2-1 Scatterplot of Rise Time threshold by Word Score (AMC and DLD groups)

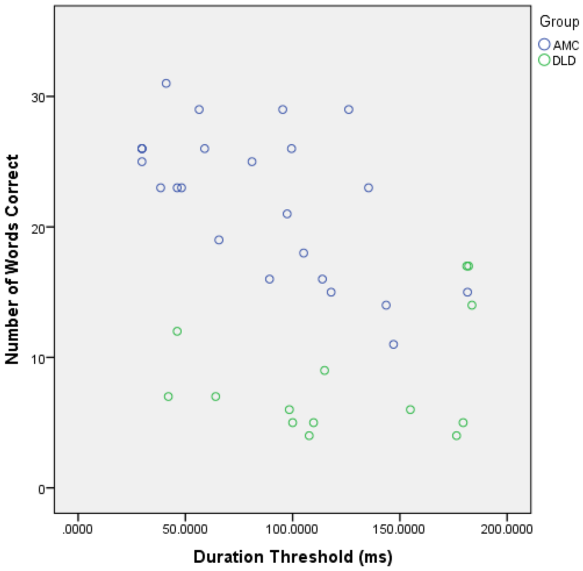


Figure B-2-2 Scatterplot of Duration threshold by Word Score (AMC and DLD groups)

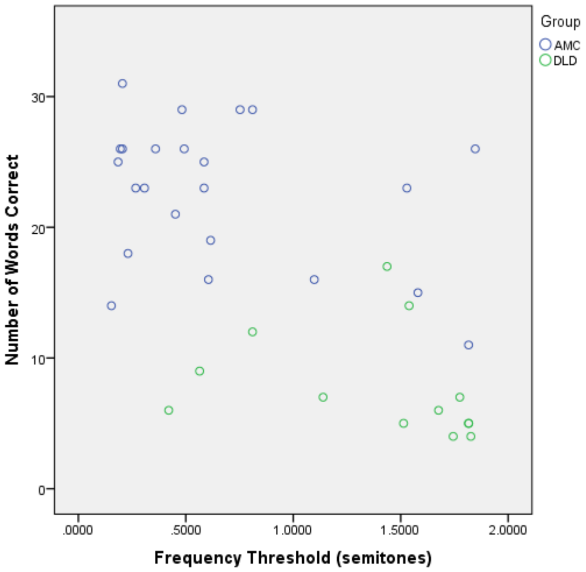


Figure B-2-3 Scatterplot of Frequency threshold by Word Score (AMC and DLD groups)

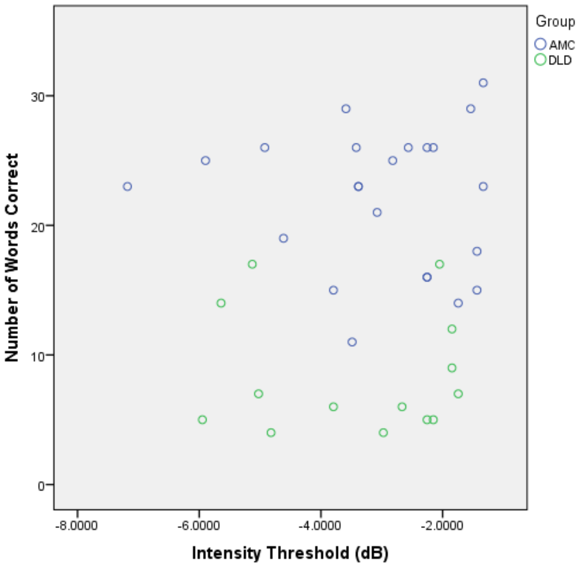
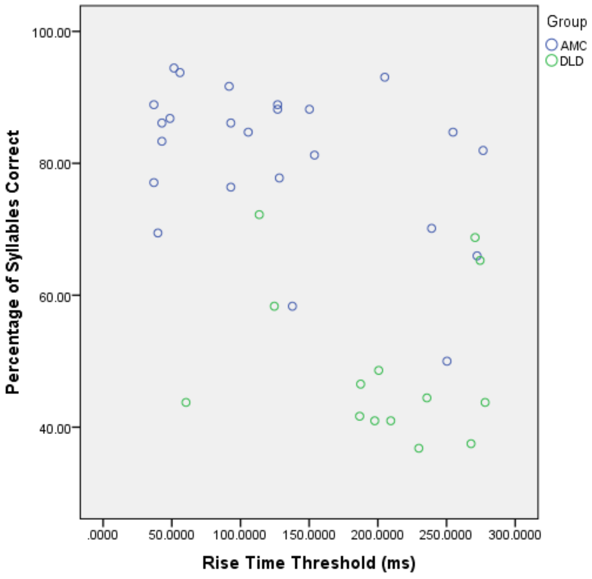


Figure B-2-4 Scatterplot of Intensity threshold by Word Score (AMC and DLD groups)

ii) Syllable Score



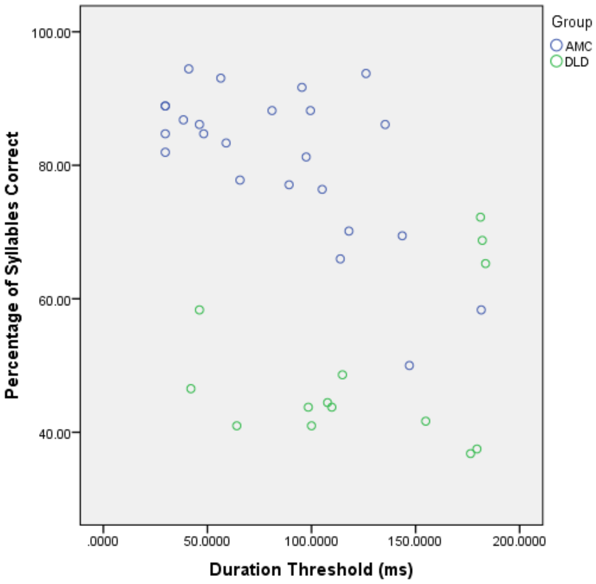
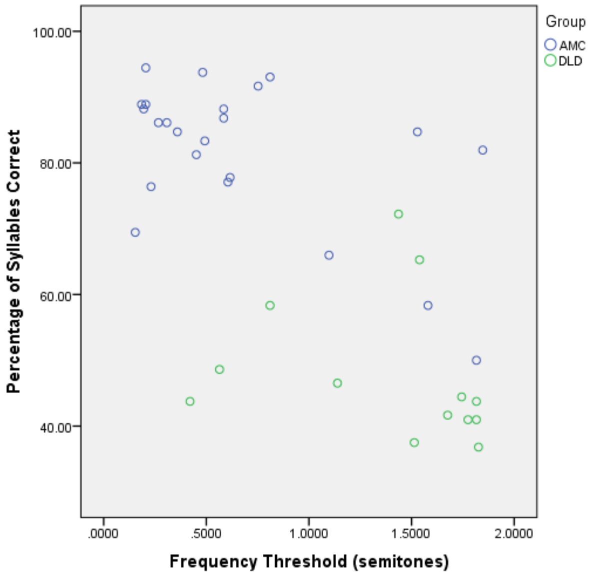


Figure B-2-5 Scatterplot of Rise Time threshold by Syllable Score (AMC and DLD groups)

Figure B-2-6 Scatterplot of Duration threshold by Syllable Score (AMC and DLD groups)



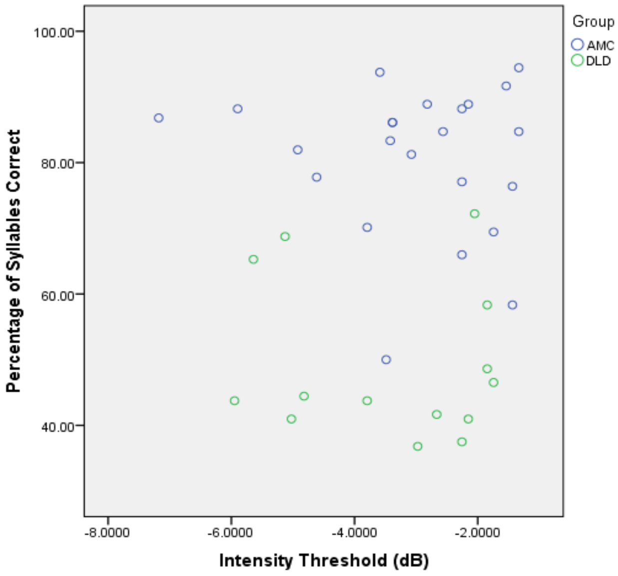


Figure B-2-7 Scatterplot of Frequency threshold by Syllable Score (AMC and DLD groups)

Figure B-2-8 Scatterplot of Intensity threshold by Syllable Score (AMC and DLD groups)

3. Experimental Task 3 – Lexical Stress

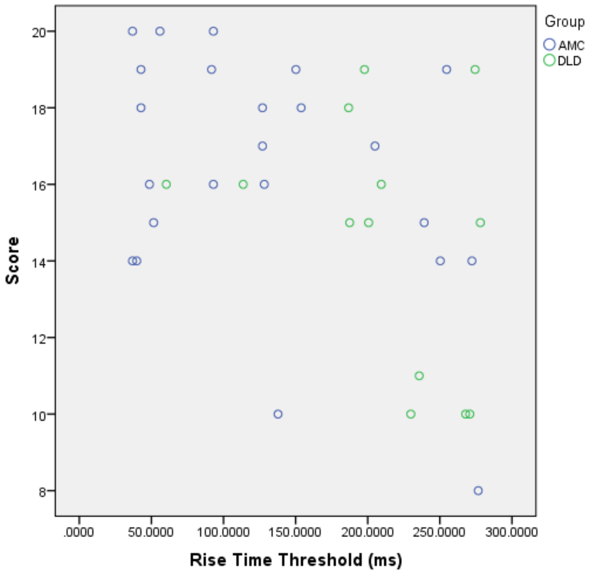


Figure B-3-1 Scatterplot of Rise Time threshold by Score (AMC and DLD groups)

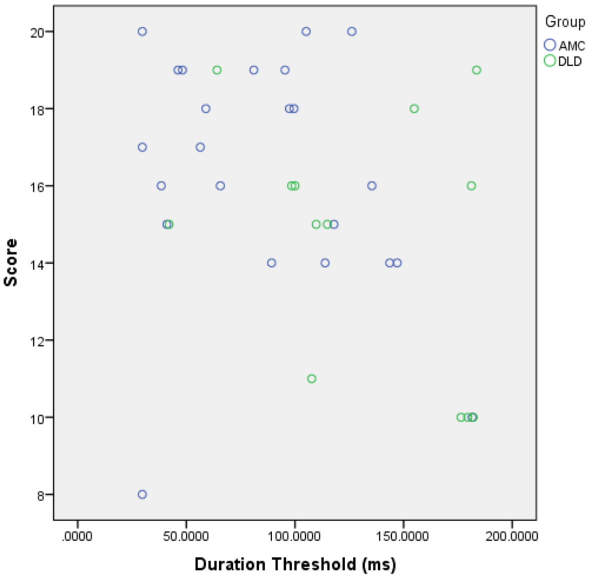


Figure B-3-2 Scatterplot of Duration threshold by Score (AMC and DLD groups)

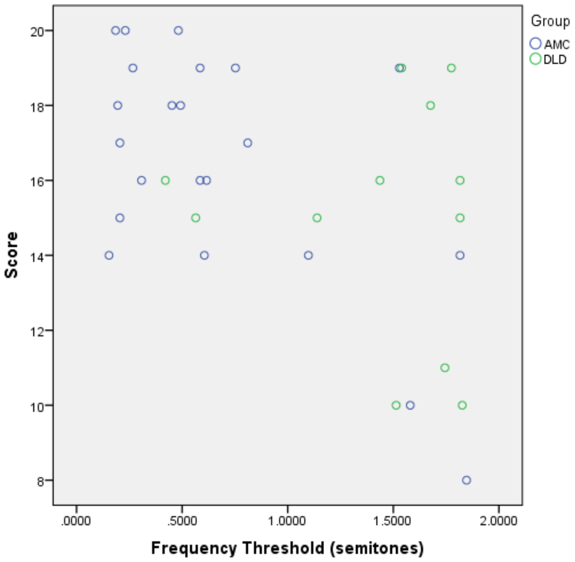


Figure B-3-3 Scatterplot of Frequency threshold by Score (AMC and DLD groups)

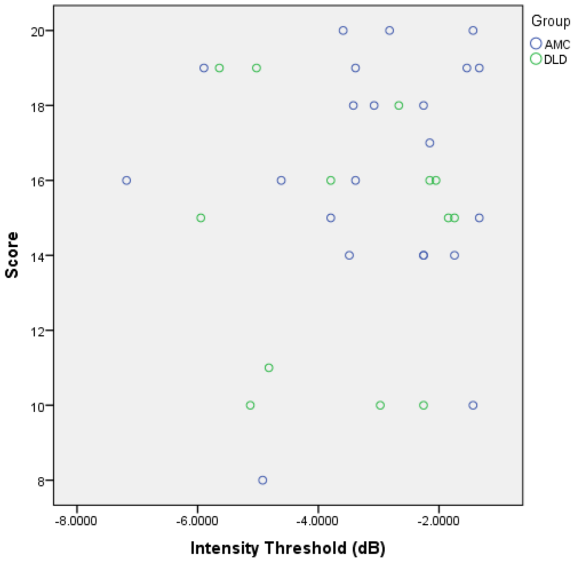
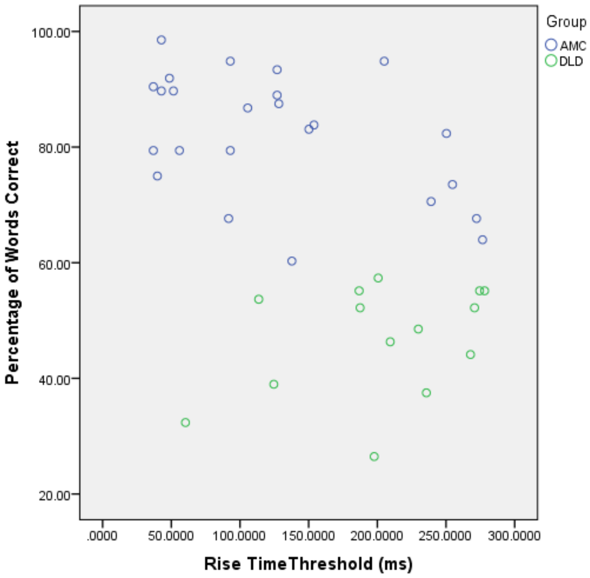


Figure B-3-4 Scatterplot of Intensity threshold by Score (AMC and DLD groups)

4. Experimental Task 4 – Sentence Repetition

i) Word Score



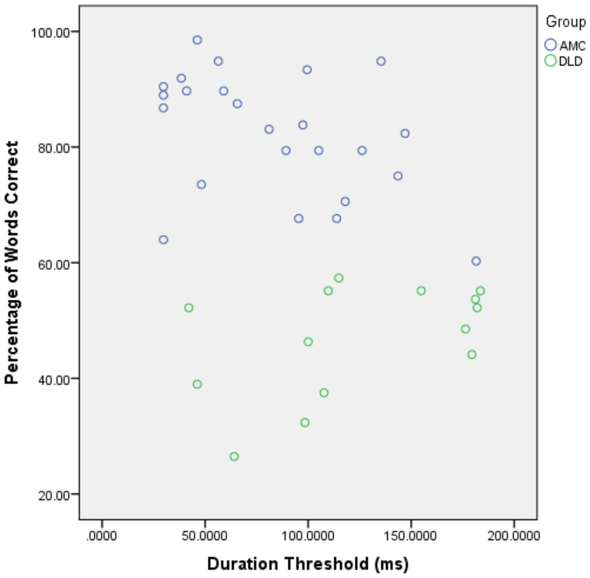
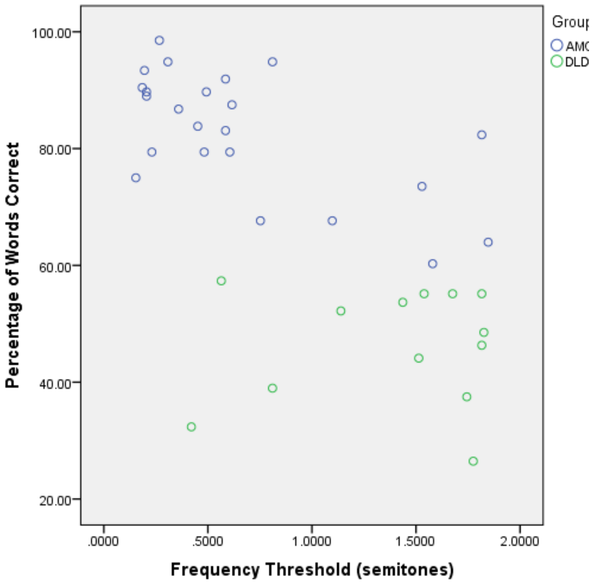


Figure B-4-1 Scatterplot of Rise Time threshold by Word Score (AMC and DLD groups)

Figure B-4-2 Scatterplot of Duration threshold by Word Score (AMC and DLD groups)



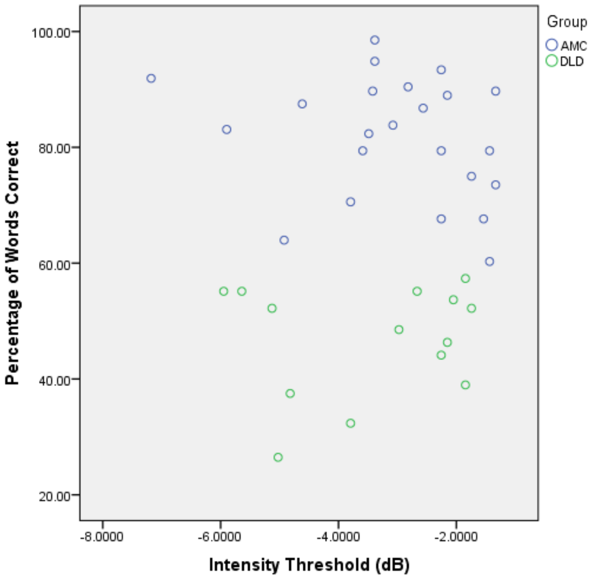


Figure B-4-3 Scatterplot of Frequency threshold by Word Score (AMC and DLD groups)

Figure B-4-4 Scatterplot of Intensity threshold by Word Score (AMC and DLD groups)

ii) Syllable Score

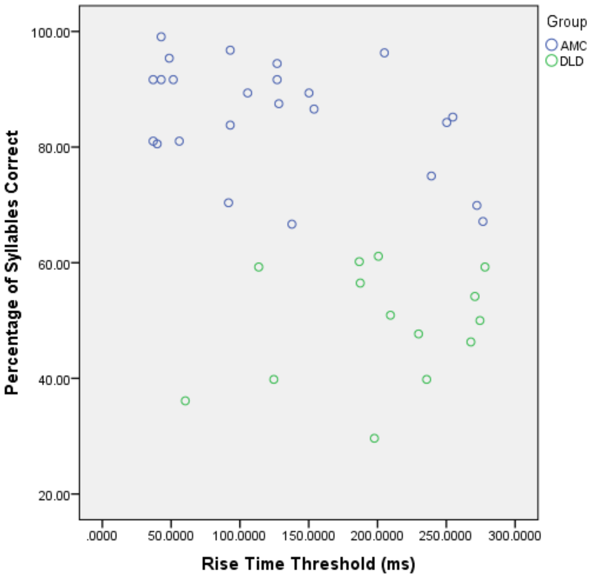


Figure B-4-5 Scatterplot of Rise Time threshold by Syllable Score (AMC and DLD groups)

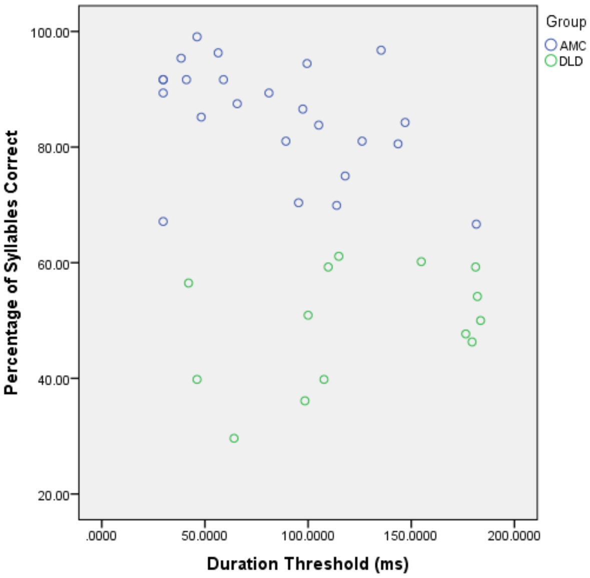


Figure B-4-6 Scatterplot of Duration threshold by Syllable Score (AMC and DLD groups)

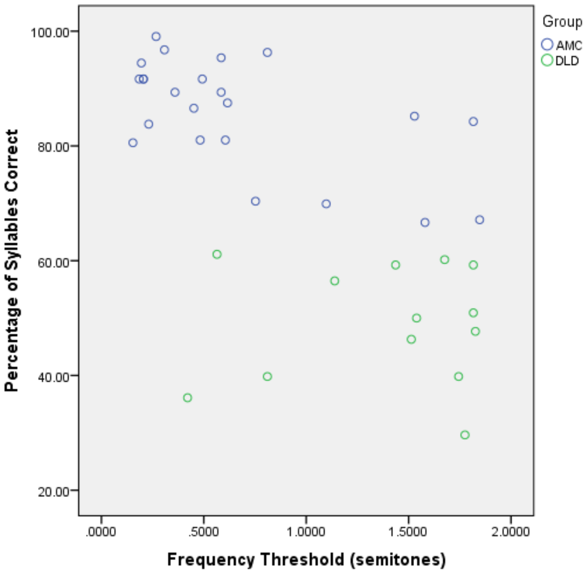


Figure B-4-7 Scatterplot of Frequency threshold by Syllable Score (AMC and DLD groups)

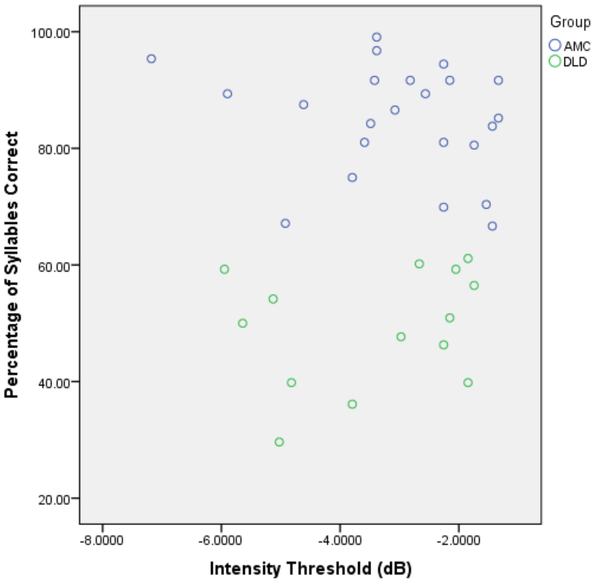


Figure B-4-8 Scatterplot of Intensity threshold by Syllable Score (AMC and DLD groups)

5. Experimental Task 5 – Metrical Stress

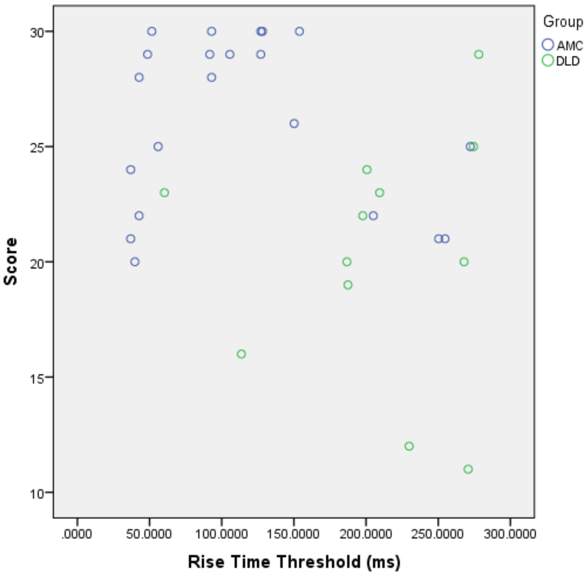


Figure B-5-1 Scatterplot of Rise Time threshold by Score (AMC and DLD groups)

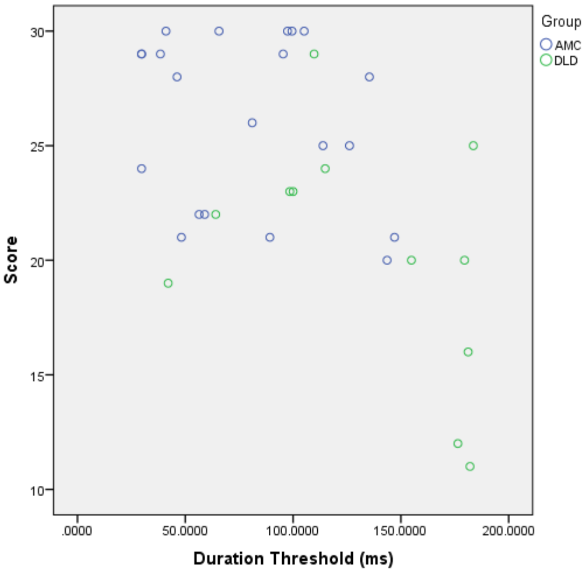


Figure B-5-2 Scatterplot of Duration threshold by Score (AMC and DLD groups)

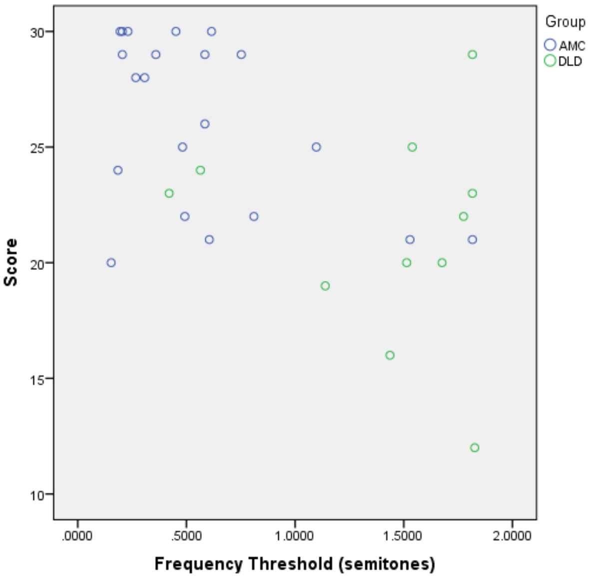


Figure B-5-3 Scatterplot of Frequency threshold by Score (AMC and DLD groups)

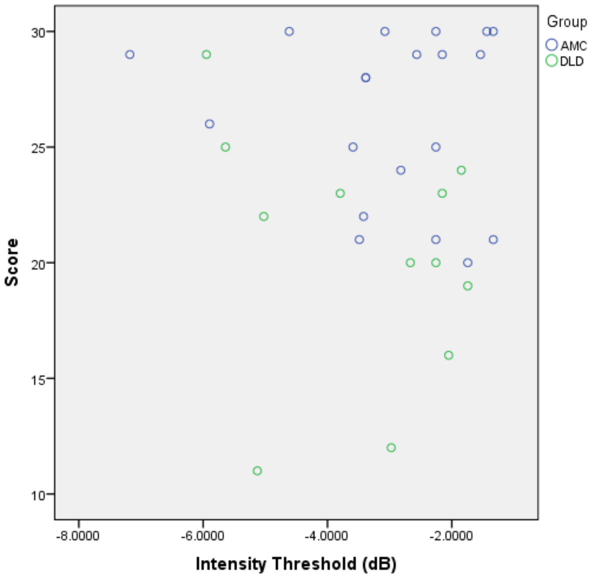


Figure B-5-4 Scatterplot of Intensity threshold by Score (AMC and DLD groups)

Appendix C – Stimuli used in the Experimental Tasks

1. Experiment 1 - Lexical Boundaries

List 1

Table C-0-1-1 Lexical Boundaries stimuli List 1

	Target Items			Number of Items
1		spaceships	and trees	2
2	snow	men	and spades	3
3	rain	bows	and biscuits	3
4		goldfish	and scissors	2
5		snowballs	and pencils	2
6		buttercups	and vans	2
7	dough	nuts	and butter	3
8		ice cream	and tables	2
9	post	boxes	and pens	3
10	jelly	fish	and books	3
11	lunch	boxes	and cars	3
12		hairbrushes	and juice	2

List 2

Table C-1--0-2 Lexical Boundaries Stimuli List 2

	Target Items			**Number of Items**
1		postboxes	and pens	2
2	hair	brushes	and juice	3
3		doughnuts	and butter	2
4		lunchboxes	and cars	2
5	gold	fish	and scissors	3
6		jellyfish	and books	2
7	space	ships	and trees	3
8	snow	balls	and pencils	3
9	butter	cups	and vans	3
10	ice	cream	and tables	3
11		snowmen	and spades	2
12		rainbows	and biscuits	2

2. Experiment 2 - Nonword Repetition

List 1

Table C-2-1 Nonword Repetition Stimuli List 1. Shading indicates stressed syllable.

	Syllables					
	1	**2**	**3**	**4**	**5**	**Wordlength: Stress**
1	kaɪ	gi	nu	fəɪ		4:1
2	fəʊ	di	kaɪ			3:2
3	pəʊ	ni	ku			3:1
4	gi	kaɪ	fəɪ	nu		4:3
5	təɪ	ku	fu	ni	pəʊ	5:4
6	pəʊ	təɪ	fu	ni		4:3
7	di	fəʊ	kaɪ	nu		4:2
8	pəʊ	təɪ	ni	ku	fu	5:2
9	di	kaɪ	fəʊ			3:3
10	təɪ	pəʊ	ni	fu		4:1
11	ni	ku	pəʊ			3:2
12	fu	ni	pəʊ	təɪ		4:2
13	fu	təɪ	naɪ			3:1
14	ku	ni	pəʊ	fu	təɪ	5:3
15	gi	nu	fəɪ			3:1
16	naɪ	fu	təɪ			3:3
17	naɪ	təɪ	ku	di		4:2
18	ku	pəʊ	ni			3:3
19	təɪ	naɪ	fu			3:2
20	fəɪ	nu	gi	kaɪ		4:2
21	təɪ	fəʊ	di	naɪ	ku	5:3
22	fəɪ	gi	nu			3:3
23	fəʊ	gi	di	kaɪ	nu	5:4
24	ku	di	naɪ	təɪ		4:3
25	di	ku	fəʊ	təɪ	naɪ	5:2
26	gi	pəʊ	fəɪ	nu	kaɪ	5:4
27	nu	fəʊ	kaɪ	gi	di	5:2
28	nu	kaɪ	fəʊ	di		4:1
29	kaɪ	nu	di	fəʊ		4:3
30	di	ku	təɪ	naɪ		4:1
31	ku	təɪ	naɪ	fəʊ	di	5:4
32	gi	kaɪ	nu	di	fəʊ	5:3
33	nu	fəɪ	gi			3:2
34	kaɪ	gi	nu	pəʊ	fəɪ	5:2
35	pəʊ	nu	kaɪ	fəɪ	gi	5:3
36	kaɪ	fəʊ	di			3:1

List 2

Table C-2-2 Nonword Repetition Stimuli List 2. Shading indicates stressed syllable.

	Syllables					
	1	**2**	**3**	**4**	**5**	**Wordlength: Stress**
1	fəɪ	kaɪ	nu	gi		4:2
2	naɪ	təɪ	fu			3:3
3	gi	nu	fəʊ	di	kaɪ	5:4
4	di	fəʊ	kaɪ			3:3
5	di	naɪ	ku	təɪ		4:3
6	nu	gi	pəʊ	kaɪ	fəɪ	5:3
7	nu	gi	fəɪ	kaɪ		4:1
8	nu	di	kaɪ	fəʊ		4:3
9	kaɪ	fəʊ	gi	nu	di	5:3
10	təɪ	fu	naɪ			3:1
11	ni	pəʊ	fu	təɪ		4:1
12	pəʊ	ni	fu	ku	təɪ	5:2
13	ni	təɪ	ku	pəʊ	fu	5:3
14	gi	fəɪ	nu			3:2
15	pəʊ	kaɪ	gi	fəɪ	nu	5:4
16	fu	naɪ	təɪ			3:2
17	ni	pəʊ	ku			3:1
18	fəʊ	kaɪ	di			3:1
19	təɪ	ku	naɪ	di		4:1
20	pəʊ	ku	ni			3:2
21	kaɪ	di	fəʊ			3:2
22	ku	pəʊ	təɪ	fu	ni	5:4
23	nu	gi	fəɪ			3:1
24	ku	ni	pəʊ			3:3
25	təɪ	fu	pəʊ	ni		4:3
26	kaɪ	nu	fəɪ	pəʊ	gi	5:2
27	fəʊ	kaɪ	di	nu		4:1
28	nu	kaɪ	di	gi	fəʊ	5:2
29	di	nu	fəʊ	kaɪ		4:2
30	naɪ	di	təɪ	ku		4:2
31	təɪ	di	ku	naɪ	fəʊ	5:4
32	di	fəʊ	naɪ	təɪ	ku	5:2
33	kaɪ	fəɪ	gi	nu		4:3
34	fəʊ	ku	təɪ	di	naɪ	5:3
35	fu	təɪ	ni	pəʊ		4:2
36	fəɪ	nu	gi			3:3

3. Experiment 3 - Lexical Stress

List 1

Table C-3-1 Lexical Stress task Stimuli List 1

	Condition	CarrierPhrase	Target
1	CorrStress-InCon	The boy is reading books about	PIllows
2	InCorrStress-Con	The boy is reading books about	coFFEE
3	CorrStress-InCon	The boy is reading books about	HOney
4	CorrStress –Con	Jack is reading books about a	BLANket
5	InCorrStress-Con	The boy is reading books about	piLLOWS
6	IncorrStress-InCon	Jack is reading books about a	peNNY
7	IncorrStress-InCon	Jack is reading books about some	hoNEY
8	InCorrStress-Con	The boy is reading books about	mounTAINS
9	CorrStress –Con	Jack is reading books about a	FOOTball
10	CorrStress-InCon	The boy is reading books about	BEDrooms
11	InCorrStress-Con	The boy is reading books about	tiCKETS
12	InCorrStress-Con	The boy is reading books about	bedROOMS
13	CorrStress-InCon	The boy is reading books about	TICKets
14	CorrStress –Con	Jack is reading books about a	PEnny
15	IncorrStress-InCon	Jack is reading books about a	chiCKEN
16	IncorrStress-InCon	Jack is reading books about a	blanKET
17	CorrStress–Con	Jack is reading books about some	COffee
18	CorrStress-InCon	The boy is reading books about	MOUNtains
19	CorrStress –Con	Jack is reading books about a	CHIcken
20	IncorrStress-InCon	Jack is reading books about a	footBALL

List 2

Table C-3-2 Lexical Stress task Stimuli List 2

	Condition	CarrierPhrase	Target
1	IncorrStress-InCon	Jack is reading books about a	bedROOM
2	CorrStress-InCon	The boy is reading books about	CHIckens
3	CorrStress-InCon	The boy is reading books about	PEnnies
4	CorrStress-InCon	The boy is reading books about	COffee
5	InCorrStress-Con	The boy is reading books about	hoNEY
6	InCorrStress-Con	The boy is reading books about	footBALLS
7	InCorrStress-Con	The boy is reading books about	peNNIES
8	IncorrStress-InCon	Jack is reading books about a	piLLOW
9	CorrStress–Con	Jack is reading books about a	MOUNtain
10	InCorrStress-Con	The boy is reading books about	blanKETS
11	CorrStress–Con	Jack is reading books about a	BEDroom
12	IncorrStress-InCon	Jack is reading books about a	mounTAIN
13	CorrStress–Con	Jack is reading books about some	HOney
14	CorrStress–Con	Jack is reading books about a	TIcket
15	CorrStress–Con	Jack is reading books about a	PIllow
16	CorrStress-InCon	The boy is reading books about	FOOTballs
17	InCorrStress-Con	The boy is reading books about	chiCKENS
18	IncorrStress-InCon	Jack is reading books about some	coFFEE
19	CorrStress-InCon	The boy is reading books about	BLANkets
20	IncorrStress-InCon	Jack is reading books about a	tiCKET

4. Experiment 4 - Sentence Repetition

List 1

Table C-4-1 Sentence Repetition Stimuli List 1

	Sentence Type	Rhythm	Dominant Prosody	Prosodic Context (Function Words)	Sentence
1	Embedded	Variable	nonSw	Sw	The awards the referee is between are real
2	Embedded	Regular	nonSw	Sw	The display the guitar is behind is complete
3	Embedded	Regular	Sw	ww	The programmes the meeting is after are awful
4	Complex	Variable	nonSw	Sw	The team were celebrating before the match had concluded
5	Embedded	Regular	nonSw	Sw	The designs the balloon is between are correct
6	Complex	Variable	Sw	ww	The audience were smiling after the performer had laughed
7	Complex	Regular	Sw	ww	The strangers were whispering after the princess had promised
8	Complex	Variable	nonSw	Sw	The crowd was disappearing before the judge had decided
9	Embedded	Variable	Sw	ww	The tables the computer is under are big
10	Embedded	Regular	Sw	ww	The carpet the metal is under is purple
11	Complex	Regular	Sw	ww	The writer was practising after the author had published
12	Complex	Variable	Sw	ww	The computer was waiting after the equipment had failed
13	Embedded	Variable	nonSw	Sw	The machine the magazine is behind is short
14	Complex	Regular	nonSw	Sw	The machines were attacking before the defeat had begun
15	Complex	Regular	nonSw	Sw	The hotel was collapsing before the police had prepared
16	Embedded	Variable	Sw	ww	The package the recorder is under is white

List 2

Table C-4-2 Sentence Repetition Stimuli List 2

	Sentence Type	Rhythm	Dominant Prosody	Prosodic Context (Function Words)	Sentence
1	Embedded	Regular	Sw	ww	The meetings the practice is after are useful
2	Embedded	Variable	Sw	ww	The carpets the tomato is under are red
3	Complex	Regular	nonSw	Sw	The reward was decreasing before the mistakes had improved
4	Complex	Regular	Sw	ww	The soldier was travelling after the battle had finished
5	Complex	Regular	nonSw	Sw	The machines were attacking before the defeat had begun
6	Complex	Variable	nonSw	Sw	The class were interrupting before the reply had finished
7	Embedded	Regular	nonSw	Sw	The reports the award is between are superb
8	Embedded	Variable	nonSw	Sw	The guitar the cigarette is behind is big
9	Complex	Variable	Sw	ww	The supporters were crying after the player had fallen
10	Embedded	Regular	nonSw	Sw	The machine the police are behind is unique
11	Embedded	Variable	nonSw	Sw	The balloons the engineer is between are green
12	Complex	Regular	Sw	ww	The students were arguing after the meeting had started
13	Complex	Variable	Sw	ww	The explorer was resting after the adventure had stopped
14	Embedded	Regular	Sw	ww	The picture the spiders are under is perfect
15	Complex	Variable	nonSw	Sw	The guard was investigating before the crime had happened
16	Embedded	Variable	Sw	ww	The pocket the potato is under is blue

5. Experiment 5 - Metrical Stress

List 1

Table C-5-1 Metrical Stress Stimuli List 1

Text removed because copyright clearance has not been granted. Copyright holder is Julia Donaldson.

Text removed because copyright clearance has not been granted. Copyright holder is Julia Donaldson.

List 2

Table C-5-2 Metrical Stress Stimuli List 2

Text removed because copyright clearance has not been granted. Copyright holder is Julia Donaldson.

List 3

Table C-5-3 Metrical Stress Stimuli List 3

Text removed because copyright clearance has not been granted. Copyright holder is Julia Donaldson.

Text removed because copyright clearance has not been granted. Copyright holder is Julia Donaldson.

Appendix D - Scatterplots for Age and Acoustic Thresholds

1. Rise Time

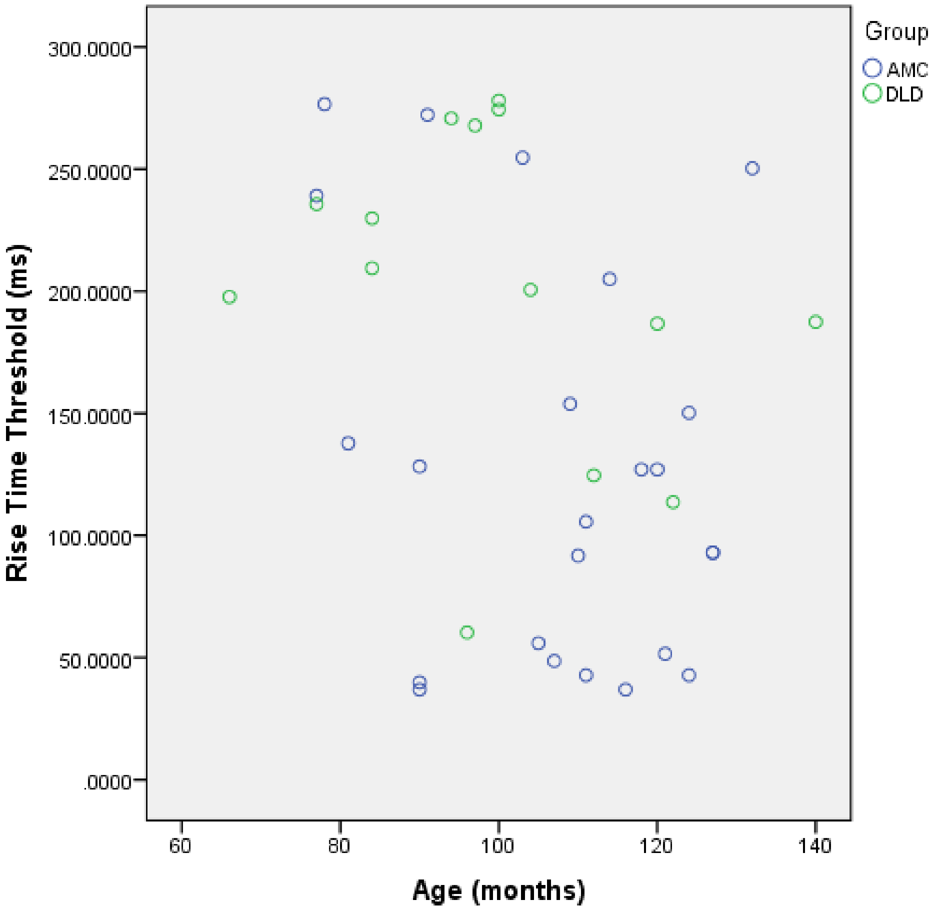


Figure D-1 Scatterplot of Age by Rise Time Threshold

2. Duration

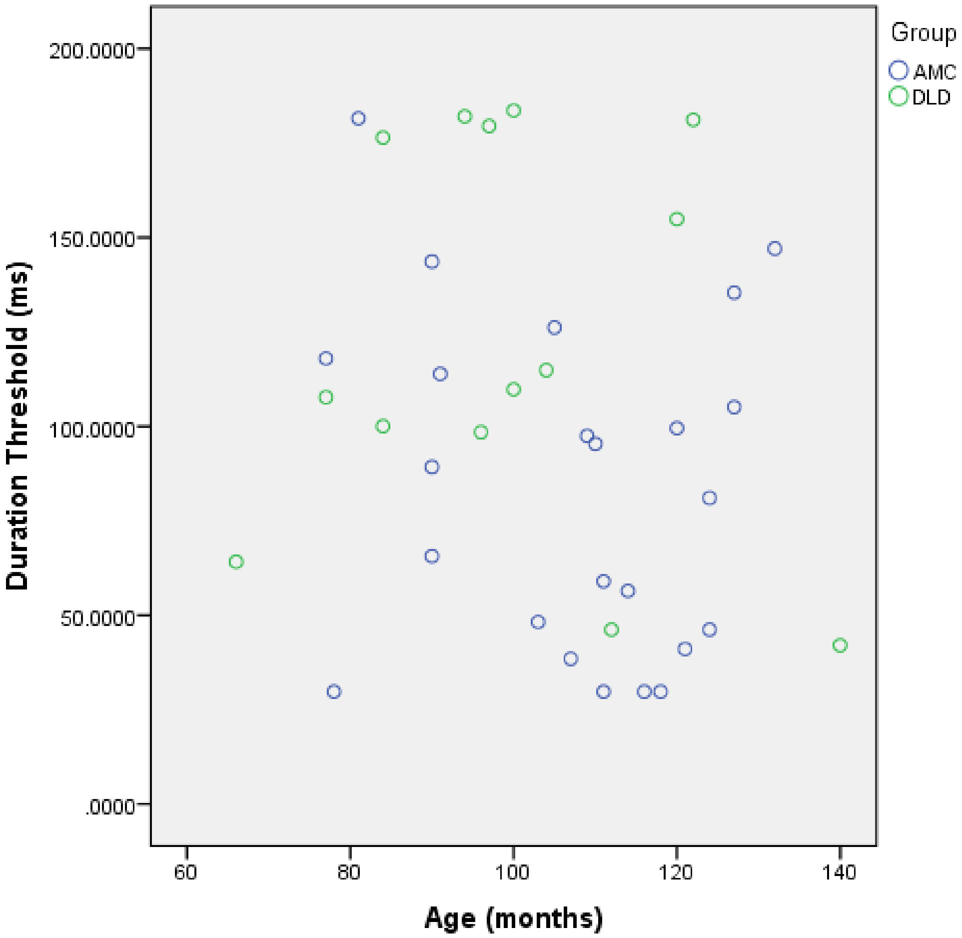


Figure D-2 Scatterplot of Age by Duration Threshold

3. Frequency

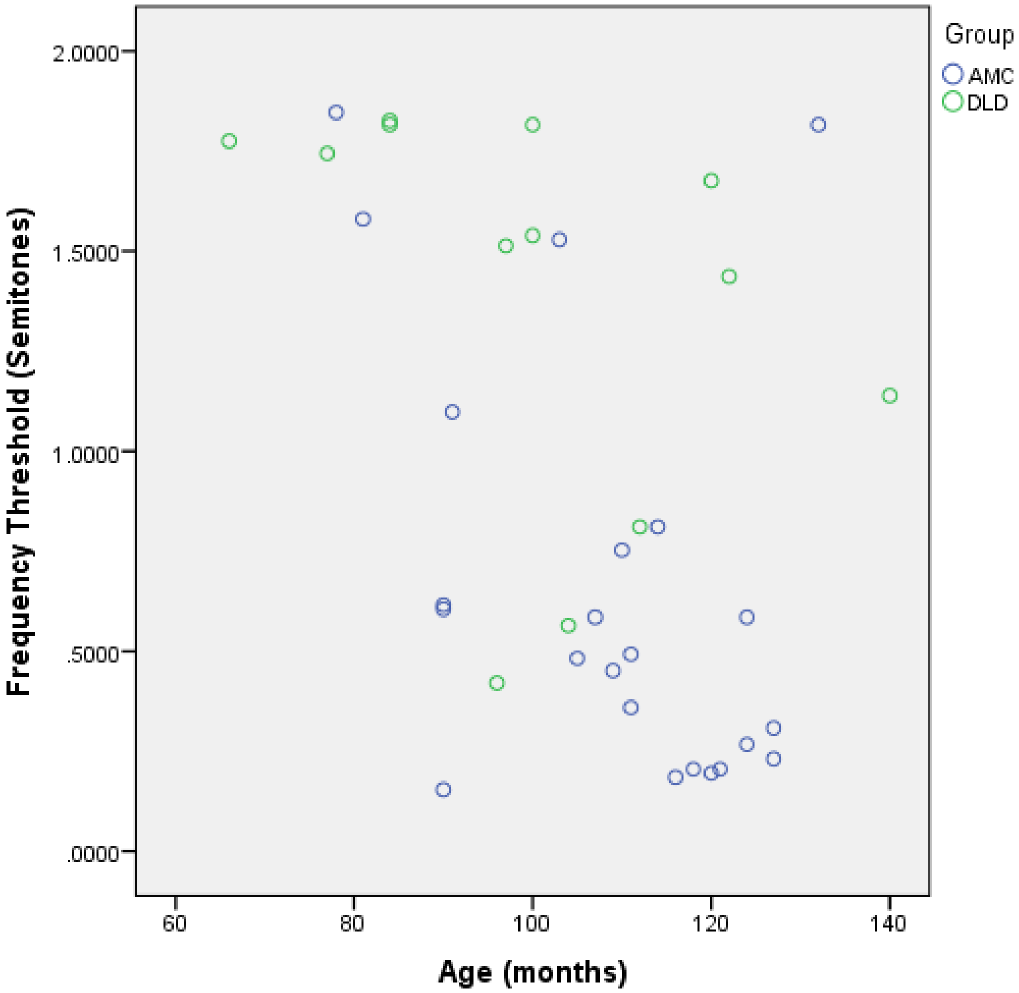


Figure D-3 Scatterplot of Age by Frequency Threshold

4. Intensity

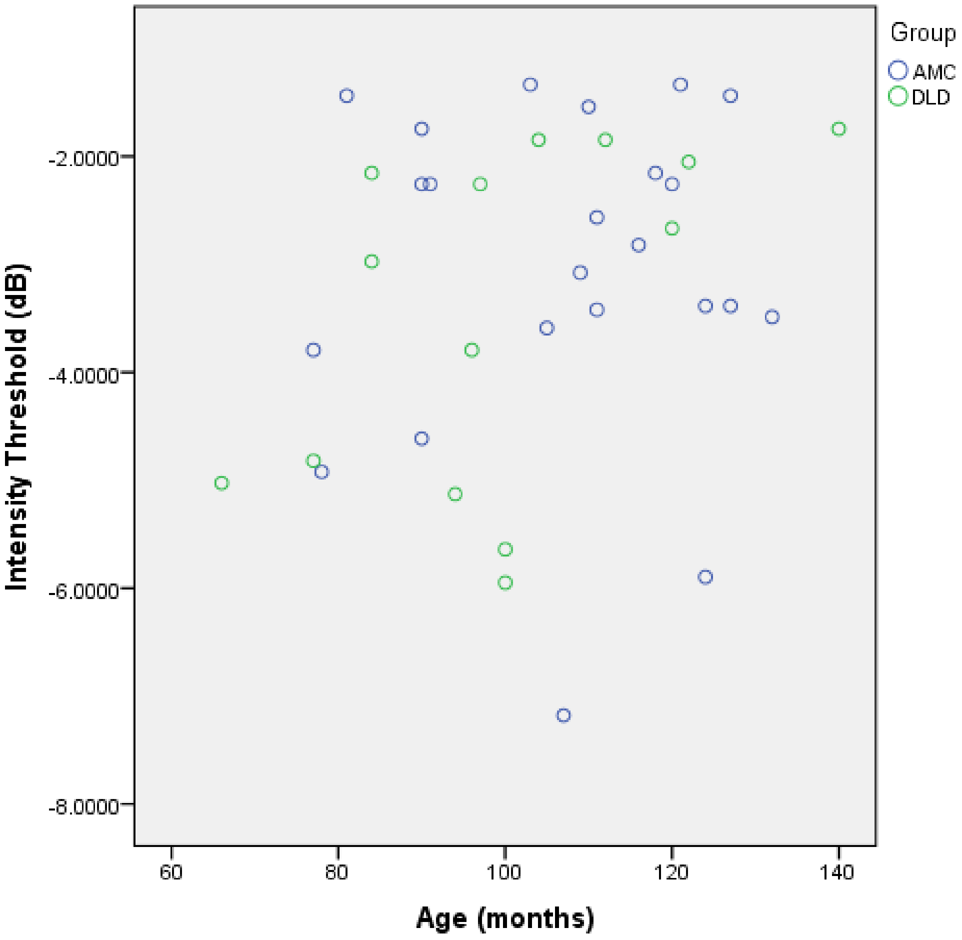


Figure D-4 Scatterplot of Age by Intensity Threshold

Appendix E - Scatterplots for Age and Task Scores

1. Experimental Task 1 - Lexical Boundaries

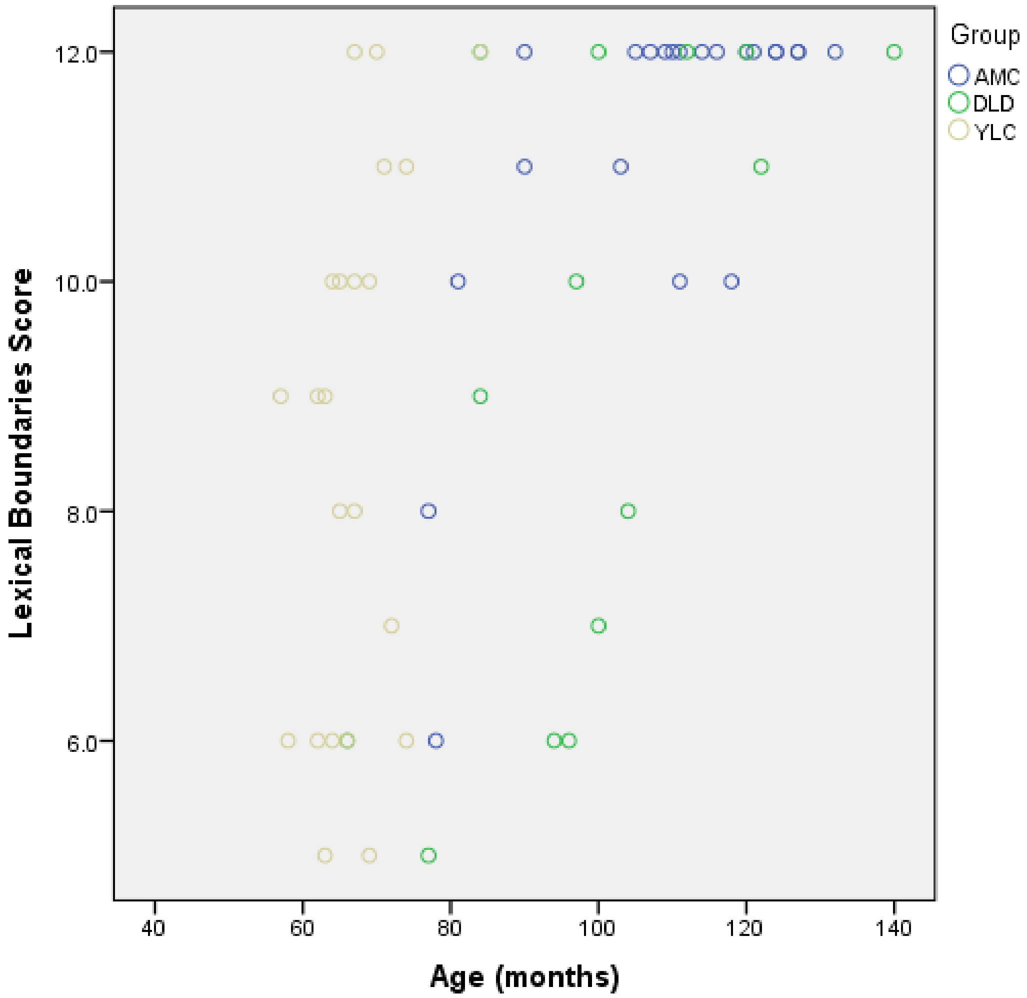


Figure E-1 Scatterplot of Score by Age (AMC, DLD and YLC Groups)

2. Experimental Task 2 – Nonword Repetition

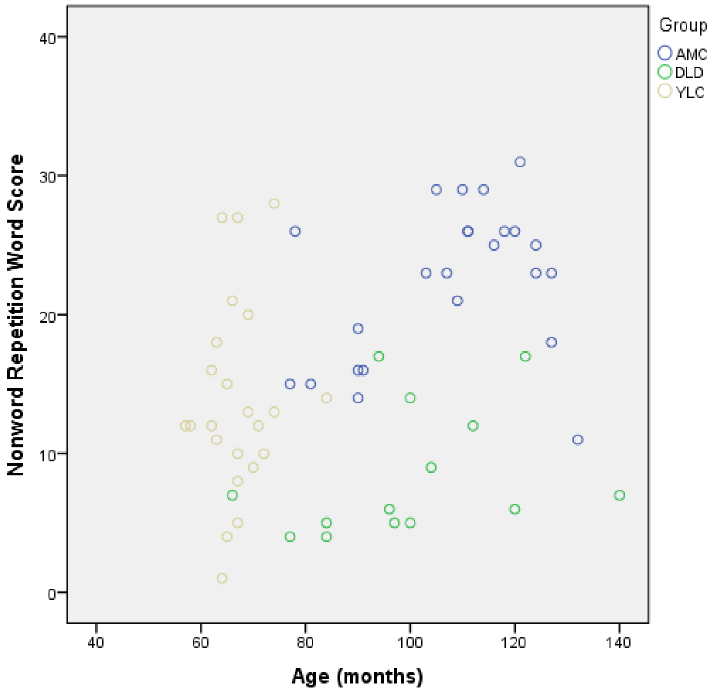


Figure E-2-1 Scatterplot of Word Score by Age (AMC, DLD and YLC Groups)

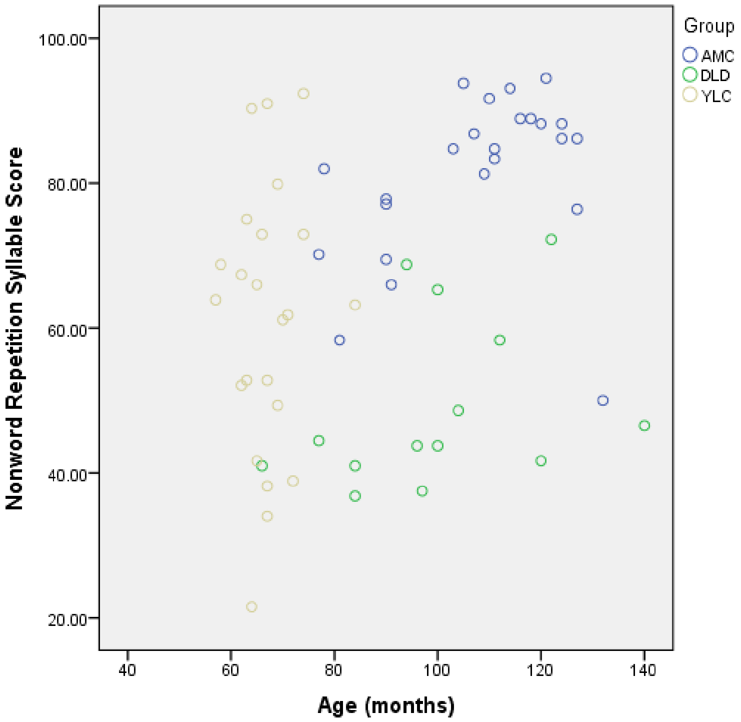


Figure E-2-2 Scatterplot of Syllable Score by Age (AMC, DLD and YLC Groups)

3. Experimental Task 3 – Lexical Stress

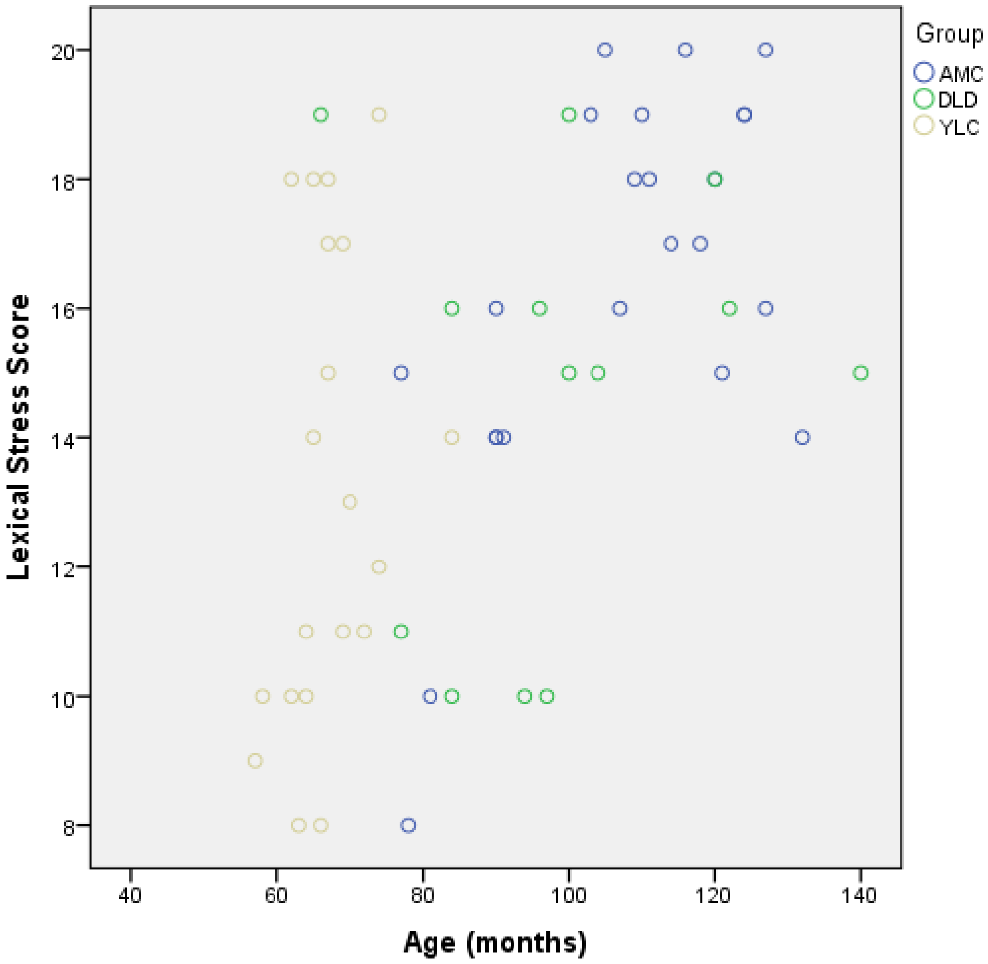


Figure E-3 Scatterplot of Score by Age (AMC, DLD and YLC Groups)

4. Experimental Task 4 – Sentence Repetition

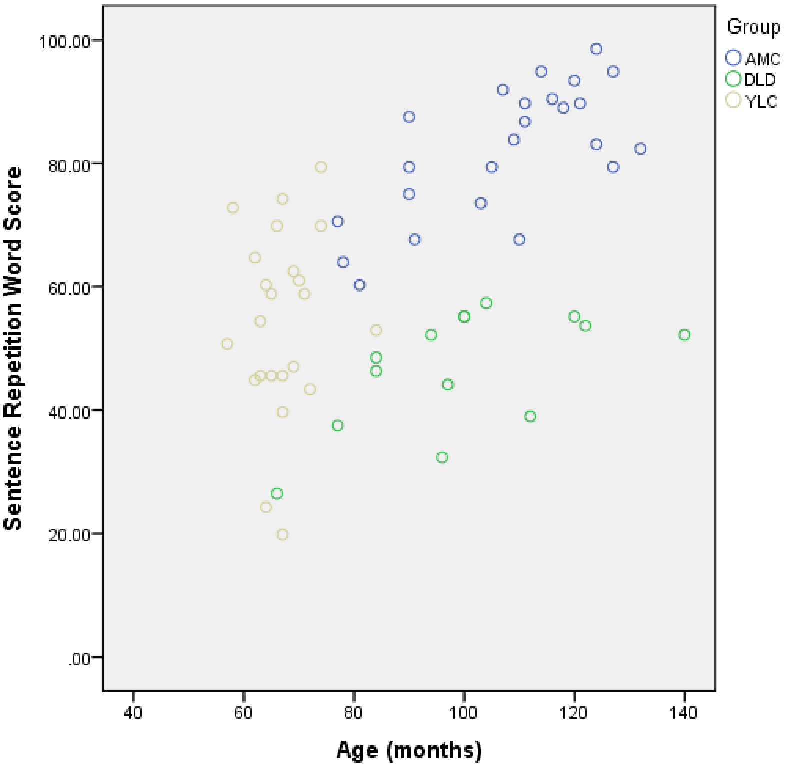


Figure E-4-1 Scatterplot of Word Score by Age (AMC, DLD and YLC Groups)

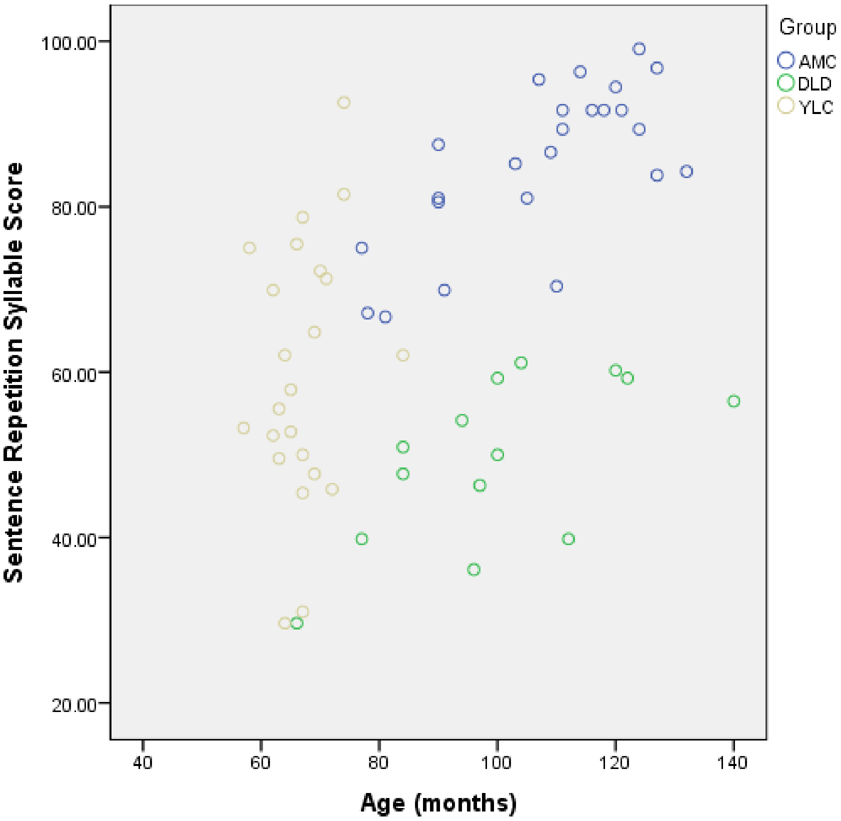


Figure E-4-2 Scatterplot of Syllable Score by Age (AMC, DLD and YLC Groups)

5. Experiment 5 – Metrical Stress

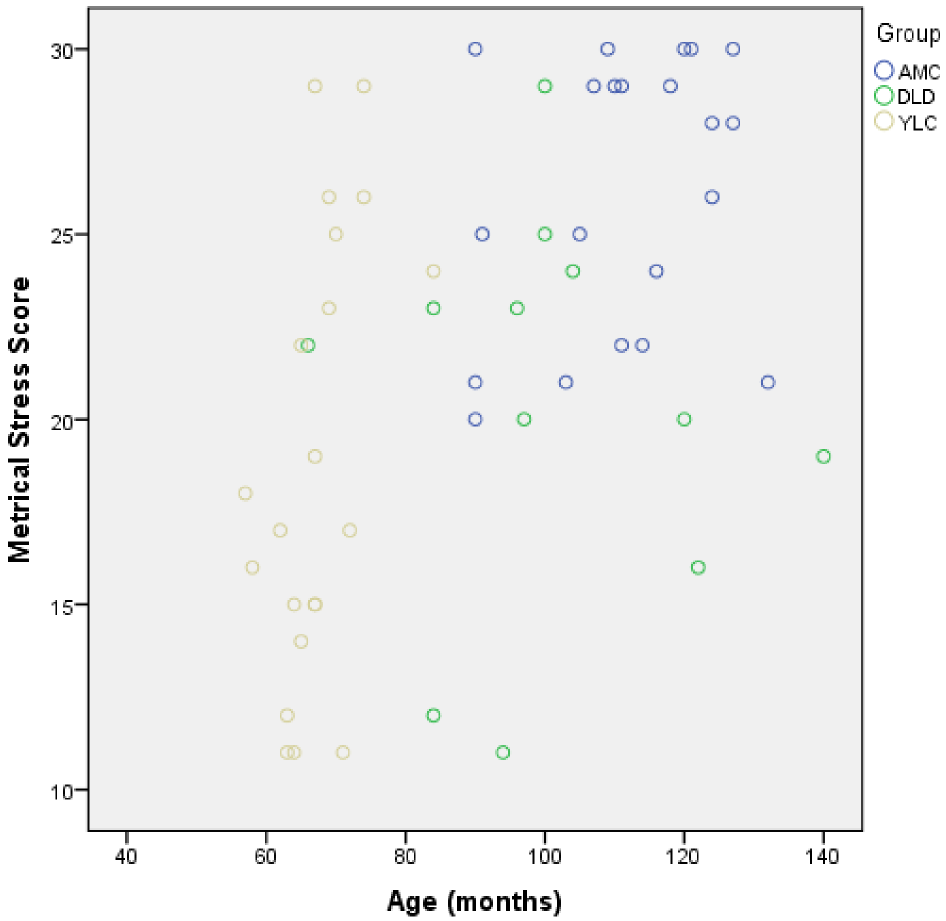


Figure E-5 Scatterplot of Score by Age (AMC, DLD and YLC Groups)