University of Strathclyde

School of Psychological Sciences and Health

Speech and Language Therapy

Rhythm in Ataxic Dysarthria

by

Jessica C. Henrich

A thesis presented in fulfilment of the requirements for the degree of Master of Philosophy

2015

Declaration of Author's Authenticity

This thesis is the results of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.

Declaration of Author's Rights

ACKNOWLEDGEMENTS

I would like to take this opportunity to thank all those people who contributed, in whatever manner, to the completion of this thesis.

I am grateful to my supervisor, Anja Lowit, for her continuous support and assistance and for not giving up on me. Thanks also to my second supervisor Ineke Mennen.

Special thanks go to the participants of the study for sharing their time and allowing me to analyse their speech and Ellika Schalling for providing speech data.

Thanks to the University of Strathclyde and the DAAD (German Academic Exchange Service) for funding me.

I also would like to thank Simone Kirstein, Caroline Bender and Anja Kuschmann for their comments, expertise, advice and continuous encouragement.

Finally, I would like to thank my parents and friends for their constant encouragement and support, with special thanks to my husband, Jörg Henrich, for his patience, love and always believing in me.

Table of Contents

List of Tables

List of figures

Abstract

Ataxic dysarthria is a speech disorder in which a disturbance of rhythm is one of the main characteristics. Although this feature has been recognised since some time ago, little detail is known on the specific manifestations of this problem. Acoustic analysis of rhythm can go some way towards extending our knowledge in this area, however, insufficient information is available on which measures are best suited to such investigations. Acoustic rhythm metrics have largely been developed for crosslinguistic comparisons of unimpaired speakers, and further research needs to be performed to identify the most suitable methodology for the evaluation of disordered speech. This study contributes to filling this knowledge gap by investigating the rhythmic differences between (cerebellar) ataxic dysarthric speech and normal speech with a range of analysis methods and speech tasks, with the aim to identify which acoustic rhythm metric is most suited to differentiate disordered from healthy speech performance, which task(s) can highlight such differences most, and which rhythm metrics correlates best with perceptual evaluations.

Different speech samples, ranging from structured to unstructured, from six speakers with ataxic dysarthria and six age and gender matched control speakers were analysed with five different rhythm measures: the normalized Pairwise Variability Index (nPVI), the VarcoV, the Proportion of Vocalic Intervals (%V), the Variability Index (VI), and the Interstress Interval measure (ISI). The results of the acoustic analysis were compared with a perceptual evaluation of the participants' speech.

Results varied between different rhythm metrics and speech tasks, but nPVI and VarcoV seem to be the metrics most suitable to characterize rhythmic changes in ataxic dysarthria. These two yielded significant differences between the ataxic dysarthric and control group, and furthermore correlated significantly with the perceptual evaluation of rhythm. However, the VarcoV and nPVI (as well as the %V) metric also correlated significantly with articulation rate and future research will need to investigate further the impact this can have on rhythmic analysis in disordered speakers. In relation to task choice, the results indicate that spontaneous speech samples are a suitable task to highlight rhythmic disturbances.

"For all its ordinariness,
speech is a remarkable and unique motor accomplishment."

(Kent, 2000:392)

I. Introduction

Speech is one of the most important means of communication. Through speech, a diversity of emotional and informative messages is transferred. Disturbances in the field of speech and communication can have extensive influences on all areas of life of the affected person. This may lead to the avoidance of certain communicative situations to the point of the isolation of a person with a speech disorder. Therefore, it is important to investigate speech disturbances to find out more about the disorders and keep the negative effects for the affected person as minimal as possible or help to find better ways to deal with a speech disturbance.

This study focuses on (cerebellar) ataxic dysarthria. It is often reported in the literature that individuals with ataxic dysarthria place stress on every syllable. This phenomenon is often referred to as "scanning speech" due to the perceptual impression of syllable-by-syllable articulation (Ackermann & Hertrich, 1994; Duffy 2012, Hartelius, Runmaker, Andersen, & Nord 2000). The perceived rhythm therefore differs from the speech rhythm of healthy control speakers. English speech rhythm (a stress-timed rhythm, Dauer 1983) is characterized by an alternation of stressed and unstressed syllables. This alternation is said to be affected in ataxic dysarthria, so that syllables appear more equal in length (Ackermann & Hetrich 1994, Duffy 2012). Therefore, the rhythm of ataxic dysarthric speech has been characterized as syllable-timed.

Investigation into ataxic dysarthric speech is often carried out by the means of perceptual methods. However, recently researchers have attempted to quantify the rhythmic disturbances through acoustic measures of phonetic units such as vowels, syllables, or stress units. A wide range of different rhythm measures is available which have been developed for different speaker groups and materials. The metrics are based on different measurements and algorithms. Several studies tested different metrics on some speech tasks on different languages in one study; and some rhythm metrics emerged as being able to differentiate different languages and speech types (amongst others: Arvaniti 2012, Wiget, White, Schuppler, Grenon, Rauch, Mattys 2010, White & Mattys 2007 a,b, Low, Grabe, Nolan 2000). Furthermore, some studies successfully tested whether the rhythm metrics are able to distinguish dysarthric from healthy control speakers (Liss, White, Mattys, Lansford, Lotto, Spitzer & Caviness 2009).

Even though metrics have shown to successfully distinguish dysarthric and healthy speech, only few studies (e.g. Lowit 2014) have performed a comparison of a wide range of measures across a variety of speech tasks and correlated results with the perceptual evaluation of the participants' speech rhythm.

The current study tested some established rhythm metrics and looked how suitable each measure is to distinguish ataxic from normal controls in different speech tasks and how comparable the results of the rhythm metrics are with each other and with a perceptual evaluation of the participants' speech rhythm.

I.I Structure of the thesis

The first chapter gives a brief overview of the background for the study and covers the areas of degenerative ataxias, ataxic dysarthria, rhythm, stress- and syllable-timing, different rhythm metrics, and perceptual evaluation of dysarthric speech. At the end of the chapter the research questions and hypotheses for this study are presented.

Chapter 2 describes the methodology for the study, followed by the results of the study which are detailed in Chapter 3. The last chapter is the discussion and conclusion in which the major findings of the study are highlighted. Furthermore, it includes the limitations of the study, as well as the implication for future research.

1. Literature Review

The following chapter has five main sections, which provide the background information for this study on rhythm in ataxic dysarthric speech.

First, the causes and nature of ataxia, in particular degenerative ataxias are described. This is followed by a definition of dysarthria and the speech symptoms of speakers with ataxic dysarthria are presented to give information for the understanding the nature of the speech symptoms. The third section provides a definition of rhythm, reviews literature on research into rhythm in ataxic dysarthria and introduces different rhythm measures. This is followed by information on perceptual evaluation of dysarthric speech. Finally, the research questions of the study are laid out.

1.1 Ataxia

The participants of this study suffer from spinocerebellar and cerebellar ataxia due to a degeneration of the cerebellum. Ataxia refers to an in-coordination or unsteadiness of posture, limbs or gait, with the speech often being involved. According to Gilman (1986, as cited in Cannito & Marquart 1997:219) ataxia is defined as the "loss of motor synergy or the ability to integrate movement subcomponents in the appropriate time and space". Specifically, it is "a disorder of the control of force and timing of movements leading to abnormalities of speed, range, rhythm, starting and stopping" (Walker 1990:365). The aforementioned clinical signs can occur together, but also in isolation and can depend on the aetiology and localisation of the disturbance within the cerebellum. Ataxia and related symptoms can be caused by different aetiologies, such as degenerative conditions, traumatic conditions, vascular conditions, infectious conditions, metabolic conditions, neoplastic conditions or congenital abnormalities, enzyme deficits or catalytic abnormalities (Cannito & Marquardt 1997, Opal & Zoghbi 2002).

The different aetiologies mentioned above may result in different symptoms to be more present than others. Even though no one to one connection between symptoms and aetiology can necessarily be found, it is assumed that patients with the same aetiology show similar symptoms. The current study focuses on participants with degenerative ataxias, to keep the subject group as homogenous as possible.

1.1.1 Degenerative Ataxias

The understanding of underlying neurological disturbances of degenerative ataxias is important for the understanding of spinocerebellar and cerebellar ataxic dysarthria (henceforth ataxic dysarthria). As there is no standardised classification of the degenerative cerebellar diseases, the current study follows the nomenclature of Klockgether, Bürk, Auburger & Dichgans (1995)[1].

Degenerative ataxia comprises a wide range of hereditary as well as adventitious diseases whose cardinal symptom is ataxia. From the

[1] The following section is based on the article by Klockgether et al. (1995) on the classification and diagnostics of degenerative ataxias, if not stated otherwise.

neuropathological point of view the degeneration of the cerebellum is the cause of those diseases.

The core function of the cerebellum is to co-ordinate muscular contractions by controlling the timing and the force of these contractions. Thus, appropriate skilled, voluntary movements can be executed in order to reach an intended aim. The cerebellum coordinates exactly timed activity through the integration of "motor output with ongoing sensory feedback" (Ghez 1991:546). It receives sensory information from the entire body and uses it to execute movements.

The degeneration of the cerebellum can occur in isolation as well as in combination with a degenerative alteration of the spinal marrow, brain stem and basal ganglia, as well as its efferent and afferent connections (Klockgether et al. 1995).

Degenerative ataxias are nowadays classified with regard to their clinical and genetic aspects; and are therefore distinguished between hereditary and adventitious ataxias. Hereditary ataxias are further distinguished in autosomal-recessive and autosomal-dominant ataxias. A third kind of hereditary ataxias are the very rare x-chromosomal inherited ataxias.

Furthermore, a distinction between adventitious ataxias and idiopathic ataxias is made, where adventitious ataxias are symptomatic ataxias with known origin and idiopathic ataxias with unknown origin (see table 1.1).

Table 1.1 Classification of Degenerative Ataxias

Hereditary Ataxias	Adventitious Ataxias
autosomal-recessive: e.g.: Friedreichs Ataxia, Vitamine E Deficiency, Ataxia teleangiectasia (Louis-Bar-Syndrome)...	Idiopathic cerebellar ataxias: pure cerebellar symptoms, e.g. Marie, Fox, Alajouanine[2]…
autosomal-dominant cerebellar ataxias: e.g.: SCA1, SCA2, SCA3 (Madacho-Joseph-Disease), ...	symptomatic ataxias: Ataxias with the following Genesis: alcohol intoxication, other intoxications, trauma…
(see also Klockgether et al. 1995)	

As the study deals with participants with spinocerebellar and cerebellar ataxia, only the autosomal-dominant cerebellar ataxias and idiophatic cerebellar ataxias are briefly described in the following section.

1.1.1.1 Autosomal-dominant ataxias

Autosomal-dominant cerebellar ataxias comprise a heterogeneous group of autosomal-dominant inherited ataxias, with different underlying combinations of the atrophy of the cerebellum, brain stem or medulla. The autosomal-dominant cerebellar ataxias are also known as spinocerebellar ataxias

[2] As with other diseases ataxias were often named after the researcher who identified or discovered the disease or a patient who suffered from the disease.

(henceforth SCA) and are caused by the degeneration of the cerebellum with involvement of the spinal cord, or its efferent and afferent connections (Durr 2010, Schöls, Bauer, Schmidt, Schulte, Riess 2004).

The SCAs are caused by a faulty gene, "a translated CAG repeat expansion mutation. Expanded CAG repeats are dynamic mutations with variable length" (Schmitz-Hübsch, Coudert et al. 2008:982).

The incidence of the SCAs is estimated 5 in 100.000 persons (Evidente Evidente, Gwinn-Hardy, Caviness & Gilman 2000). Schöls et al. (2004) report a prevalence of 3 in 100.000, but assume that this does not reflect the real occurrence of the disease, as data is restricted to a few studies in isolated geographic areas. The age of onset can differ from childhood to late adulthood, and even within one family.

The clinical criteria for the diagnosis of an autosomal-dominant cerebellar ataxia are a progressive ataxia which cannot be explained otherwise, and autosomal dominant inheritance. The autosomal-dominant cerebellar ataxias are usually classified genotypically (Evidente et al. 2000). They are caused by the presence of a faulty gene, each type by a different gene. They share the same symptoms, but a high variability of the clinical features is apparent, a high overlap of phenotype occurs, and they are often only distinguishable by genetic testing (Schöls et al. 2004). Therefore there are no particular clinical features associated with distinct genetic subtypes. 27 different gene loci are known so far and genetic testing is available for 20 of them (Fonteyn, Schmitz-Hübsch et al. 2010). The spinocerebellar ataxias are named after the gene number; SCA 1, SCA 2 and so on.

The neurological symptoms include "ataxia of gait, stance and limbs, cerebellar dysarthria, oculomotor disturbances of cerebellar and supranuclear genesis, retinopathy, optic athrophy, spasticity, extrapyramidal movement disorders, peripheral neuropathy, sphincter disturbances, cognitive impairment and epilepsy" (Schöls et al. 2004).

As the specific pathophysiological mechanisms are still under investigation there is no curative treatment for the autosomal-recessive and autosomal-dominant ataxias (Fonteyn et al. 2010).

1.1.1.2 Idiopathic cerebellar ataxias

Idiopathic cerebellar ataxias are a heterogeneous group of neurodegenerative disease with unknown aetiology, with the main symptom being a progressive ataxia starting in adulthood with intermittent occurrence. Two different forms present: A pure cerebellar form with underlying cortical atrophy, and the so called "plus form" with non-cerebellar symptoms with underlying olivopontocerebellar atrophy. Clinical diagnostics show a progressive ataxia which cannot be explained otherwise with onset after the age of 25 and a negative family anamnesis and no consanguinity of the parents. If four years after onset the clinical picture shows pure cerebellar symptoms, it will be diagnosed as a cerebellar ataxia. If non-cerebellar symptoms occur, the plus form is present. Before that a definite diagnosis cannot be made. The diagnosis of multi system atrophy, which is common for the plus form, is determined on the basis of: Progressive cerebellar ataxia and/or Parkinson's syndrome with low response to L-Dopa and onset after the age of 25, severe autonomous failure, absence of dementia, viewing paresis and areflexia, negative family anamnesis and no consanguinity of the parents. The distinction of pure idiopathic cerebellar ataxia and symptomatic ataxias is important since the latter is treatable or at least further progression can be stopped (Klockgether et al. 1995).

1.2 Dysarthria

The degeneration of the cerebellum due to spinocerebellar or cerebellar ataxias leads amongst other symptoms to dysarthria (Schöls et al. 2004). Dysarthria is defined as a motor speech disorder which results from disturbances in muscular control (Darely, Aronson, Brown 1975) and is associated with "damage to neural mechanisms that regulate speech movements" (Netsell 1984:1) or lesions of different parts of the central and peripheral nervous system (Kent, Kent, Rosenbek, Vorperian & Weisemer 1997, Netsell 1984, Ziegler, Vogel, Groene & Schröter-Morasch 1998). Dysarthria "refers to a group of speech disorders characterized by disturbances in the dimension of strength, speed, tone, steadiness, accuracy, and range of movement in the muscles of speech mechanism" (Love 1995:23). It is characterized by an inability of normal regulation of speech movements; all speech processes may become deviant, such as respiration, phonation, resonance, articulation, prosody (Love 1995, Kent, Kent, Duffy, Thomas, Weismer, Stuntebeck 2000).

Disturbances in the respiratory system can affect the breathing-type, -length and coordination, whereas disturbances in the phonatory system express themselves as affected pitch, loudness, voice quality and rigidity. Disturbances in the articulation of vowels, consonants and nasality are due to articulatory problems and prosody encompasses deviant accent, rhythm, intonation and speech rate (Ziegler et al. 1998).

Different types of dysarthria exist (e.g. spastic, hypokinetic, hyperkinetic ataxic, flaccid, unilateral upper motor neuron or mixed dysarthria). Due to the focus of this thesis, only ataxic dysarthria will be discussed in greater detail. For information on the other types of dysarthria the reader is referred to Duffy (2012).

Dysarthric symptoms occur due to paralysis, weakness, or incoordination of speech musculature and can be observed in any of the components required for fluent speech, including respiration, problems with the resonating cavity, articulation, phonation and prosody (Ziegler et al. 1998). The symptoms "can vary considerably with respect to the severity of speech disorder" (Kent, Kent, Duffy & Weismer 1998:198) as well as with the type of dysarthria. Even though not all patients diagnosed with the same type of dysarthria show the exact same pattern of speech symptoms, clusters of symptoms are found in the different dysarthrias. Therefore, similar descriptions of the symptoms of speaker with ataxic dysarthria can be found across different dialects and languages (Chenery, Ingram & Murdoch 1990, Duffy 2012, Kent et al. 1998).

The term "ataxic dysarthria" in the current study corresponds to speakers who are diagnosed with a degenerative ataxia (specifically spinocerebellar and cerebellar ataxia), which caused a dysarthria in the course of the disease.

1.2.1 Symptoms of ataxic dysarthria

Several reseachers report three clusters of co-occurring deviant perceptual dimensions as typical speech symptoms of ataxic dysarthric speech (amongst others Darley, Aronson, & Brown 1969 a,b, Kent & Netsell 1975, Kent, Netsell & Abbs 1979, Hartelius et al. 2000, Schalling, Hammarberg, Hartelius 2008, Duffy 2012). These three clusters of deviant speech dimensions are articulatory inaccuracy, phonatory-prosodic insufficiency and prosodic excess (Darley et al. 1969b) (table 1.2).

Table 1.2 Speech Symptoms of ataxic dysarthria

Articulatory inaccuracy	Phonatory-prosodic insufficiency	Prosodic excess
- Imprecise consonants - Irregular articulatory breakdowns - Distorted vowels	- Harshness - Monopitch - Monoloudness - Disturbances of pitch (flat F0) - Sudden burst-like increases of loudness - Inability to sustain vowels	- Reduced stress contrasts (excess and equal stress) - Prolonged phonemes - Prolonged intervals - Slow rate
Based on Darley FL, Aronson AE, and Brown JR: Differential diagnostic patterns of dysarthria, J Speech Hear Res 12:246, 1969b / Duffy JR, Motor Speech Disorders, 1995:155		

Even though the aforementioned features are not prominent in all speakers with ataxic dysarthria, they are seen as distinctive symptom features (Duffy 2012). Additionally, Hartelius et al. (2000) point out that two or more of the presented deviant speech dimensions are mostly apparent in ataxic dysarthria, but it is likely that there is always a predominance of one cluster. In agreement with that Kent et al. (1997) described "imprecise consonants, excess and equal stress, irregular articulatory breakdowns, distorted vowels, harsh voice, prolonged phonemes and intervals and voice tremor" (Kent et al. 1997:64) as the most common features associated with cerebellar disease.

1.2.1.1 Articulatory inaccuracy

Articulatory inaccuracy presents itself as imprecise consonants, irregular articulatory breakdowns and distorted vowels (Duffy 2012, Kent et al.1997). The occurrence of articulatory breakdowns may be promoted by the repetition of multisyllabic words.

The articulation of imprecise consonants occurs most often in consonant clusters or consonants that are more difficult to pronounce. Inconsistent and sudden lengthening of one or more syllables are the characteristics of irregular articulatory breakdowns (Brown, Darley & Aronson 1970). It is assumed these articulatory problems mirror an "inaccuracy in the direction of articulatory movements and dysrhythmia of repetitive movements" (Duffy 1995:155, Kent & Netsell 1975, Kent et al. 1979). In their acoustic study Rosen, Folker, Murdoch, Vogel, Cahill, Delatycki & Corben (2011) noticed a reduced distinction between vowels and consonants and slower formant transitions in speakers with Friedreich's ataxia as compared to control speakers.

1.2.1.2 Phonatory-prosodic insufficiency

Phonatory-prosodic insufficiency is represented by harshness, monopitch or monoloudness (Duffy 2012, Kent & Netsell 1975, Kent et al. 1979, Kent et al. 1997) and may be due to an "insufficient excursion of muscles [...] as a result of hypotonia" (Duffy 1995:155). Perceptual and acoustic studies reported disturbances of pitch (flat F_0) and sudden burst-like increases of loudness as well as the inability to sustain vowels (Cannito & Marquardt 1997, Hertrich & Ackermann 1993, Gentil 1990, Kent & Rosenbek 1982). The symptoms are said to be striking if present in a speaker with ataxic dysarthria. However, they do not seem to be the most deviant symptoms (Duffy 2012).

1.2.1.3 Prosodic excess

Prosodic excess is described as a combination of one or more of the following characteristics: excess and equal stress, prolonged phonemes, prolonged intervals and slow rate (Brown et al. 1970, Cannito & Marquardt 1997, Duffy 2012, Kent et al. 1997). It therefore relates to slowness of individual and repetitive movements, lengthening of unstressed syllables, increased and equalised syllable duration, as well as almost regular spacing between syllable nuclei. This is referred to as a tendency to syllable isochrony; the alternation of stressed and unstressed syllables is affected and the syllables appear more equal in length (e.g. Kent et al. 1997). The combination of these factors leads to the perceptual impression of a word-by-word / syllable-by-syllable articulation. This syllable-by-syllable articulation, a slowed speech performance with prolonged syllables is frequently referred to as "scanning speech" (Ackermann & Hertrich 1994, Duffy 1995, Hartelius et al. 2000, Kent & Netsell 1975, Kent & Rosenbek 1982, Kent et al. 1979, Stuntebeck 2002, Ziegler & Wessel 1996). Hartelius et al. (2000) concluded that the absence of even inter-stress intervals (regular stress beats) added to the perceptual impression of scanning speech.

Research into the acoustic nature of ataxic dysarthric speech confirmed the differences found in perceptual studies: Temporal and durational reduced variability in vowel, syllable or inter-stress interval length were found (which add to the perceptual impression of scanning speech) (e.g. Ackermann & Hertrich 1994, Boutsen, Bakker & Duffy 1997, Kent & Rosenbek 1982, Murry 1983 a,b, Yorkston, Beukelmann, Minifie & Sapir 1984). A slow speaking rate and increased syllable duration have also been experimentally verified (e.g. by Ackermann & Hertrich 1994, Gentil 1990, Kent et al. 1997, Linebaugh & Wolfe 1984, Lowit, Kuschmann & MacLeod 2010). Hartelius et al. (2000) and Kent et al. (1979) found equally stressed syllable durations, involving less reduction in vowel or consonant length compared to healthy control speakers.

An increase of the duration of unstressed vowels in read speech was found by Brown & Docherty (1995). A higher variability for Swedish speakers with ataxic dysarthria as compared to healthy control speakers with problems to keep regular stress beats was found by Schalling & Hartelius (2004).

Kent & Rosenbek (1982) investigated prosodic disturbances and the relationship to neurologic lesions and reported the following acoustic characteristics of ataxic speech to be of particular interest: "Limited variation in syllable duration […] wide and nearly regular spacing between syllable nuclei", demonstrating prosodically dissociated syllables (Kent & Rosenbek 1982:262). These characteristics give the impression "of a dissociated or decomposed syllable structure and monotony" (Hartelius et al. 2000:230). Additionally, Spencer & Rogers (2005) found changes in syllable length and reduced task complexity.

1.3 Task dependent performance

As early as the 1960s research revealed significant phonetic differences between various speech tasks in healthy (normal) speakers (Abercrombie 1965, Brown & Docherty 1995, Leuschel & Docherty 1997). Abercrombie (1965) noted that every speech category (reading aloud, monologue and spontaneous speech) had "distinguishing linguistic, including phonetic peculiarities" with the most valid differences showing at a phonetic level (Abercrombie, 1965:2). Research found differences for pause duration, location and number of pauses, variation in articulation rate as well as articulatory reduction across tasks in healthy speakers (Lowit-Leuschel & Docherty 2001). Articulatory reduction may be found more often and extreme in conversation or spontaneous speech (Kohler 1995).

Comparable to healthy speakers', differences in task performances can be found for speakers with dysarthria e.g. differences between reading aloud and spontaneous speech (Leuschel & Docherty 1997). Stuntebeck (2002) found that her speakers with ataxic dysarthria differed in their performance in different sentences and she suggested that some speech tasks are more suited to reflect the characteristics of disturbed speech rhythm in ataxic dysarthria than others. Furthermore, differences in task were found for prosodic as well as segmental features, especially for speech rate, pause placement, pause duration, mean F0, F0 variation, articulator variation and number of unstressed vowels as well as unstressed vowels duration (Kent & Rosenbeck 1982, Brown & Docherty 1995, Kohler 1995, Ziegler & Wessel 1996, Leuschel & Docherty 1997, Nishio & Niimi 2001, Lowit-Leuschel & Docherty 2001). Moreover, a scanning pattern in syllable timing, inconsistent articulatory errors and breakdowns are also reported to vary depending on the speaking task (Kent et al. 1997). Additionally, the choice of speech task may also have influence on the intelligibility of speakers with dysarthria. Spontaneous speech is usually less intelligible than read speech, which might be due to different phrasing, intonation or different use of stress (Brown & Docherty 1995).

Even though task differences are visible for healthy speakers and speakers with dysarthria, it is not clear if speakers with dysarthria are able to produce modifications in different speech tasks to the same extent as healthy control speakers can. The speech performance of speakers with dysarthria across various tasks may reflect an impairment of timing regulation where some phonetic segments can be affected more than others (Kent et al. 1979). Thus researchers assume that speaking tasks influence the way dysarthric features manifest (task-dependent performance) (Kent et al. 1979, Brown & Docherty 1995, Ziegler & Wessel 1996, Kent et al. 1997, Ackermann & Hertrich 2000, Kent et al. 2000, Lowit, Miller & Poedjianto 2003).

Task dependent performance highlight reasons for the use of naturalistic and structured speech tasks to capture as many aspects of the speech disturbance in ataxic dysarthria as possible (Lowit-Leuschel & Docherty 2001). Therefore, the "choice of sampling task [...] seems [to be] an important factor in the study of prosody in the dysarthric or any other population" (Leuschel & Docherty 1997:158).

1.4 Speech Rhythm

The aforementioned speech symptoms show that speakers with ataxic dysarthria are said to show a deviant speech rhythm as one of their core speech symptoms. The rhythm is perceived as being different in ataxic dysarthric speech as compared to the speech of healthy control speakers. To investigate the reason of this perceived difference, one needs to establish first what rhythm is.

Rhythm is understood in many different ways and is present in many different aspects of our life. As stated by Eriksson (1991), definitions of rhythm always include regularity of occurence, together with a structuring of perceptual succeeding events. Handel (1989) defines rhythm as "the perception of the grouping and ordering of elements" (Handel 1989:385). Eriksson (1991) supports the idea that rhythm is a perceptual phenomenon with a strong correlation between the character of stimulus and its resulting perception. That means that the regularity and rate of successions are a relevant factor in the perception of rhythm (Eriksson 1991), e.g. when listening to a 'tango' and a 'waltz' both are perceived as rhythmic, however differently. As rhythm can and is understood in many different ways, it is important to define the concept of rhythm as it relates to speech.

The perception of rhythm in speech is influenced by the fact that humans tend "to organize several individual speech elements into regular groups" (Zellner Keller & Keller in press:2) when they produce or perceive rhythm. On the same note Handel (1989) states that listeners arrange perceptual elements into rhythmic units. Long time-intervals between two acoustic stimuli are generally perceptually underestimated by listeners; short time-intervals are generally overestimated (Allen 1975 as cited in Zellner Keller & Keller in press). This shows that listeners usually try to rhythmically structure perceived sounds and to hear groups in a series of events, even if there is no physical division into rhythmical units and the units have a perfectly regular sequence. This phenomenon is generally referred to as "subjective rhythmization", i.e. the tendency to perceive a perfectly regular sequence of identical stimuli in groups (Eriksson 1991).

A reflection of this phenomenon in different languages is the ticking of a clock: while each swing of a pendulum has the same duration as the following it is described as coming in pairs (a binary distinction): Tick-tack (German), Tic-tac (French), ticktack (Swedish), tick-tock (English) (Eriksson 1991, Setter 2000).

The perception of identical regularly spaced stimuli can, however, be influenced by different types of accents, the duration, pitch or intensity of the accents and difference between them. Whereas durational intervals between stimuli of (more than) 1.5-2.0 seconds results in the impression the stimuli are not related, 50 milliseconds seem to be the shortest interval required for stimuli to be perceived as following one another (Handel 1989). Other influencing elements are, as mentioned above, intensity, duration, and frequency.

Elements are perceived as being grouped into units of two, with the accented element starting the group, if for one element within a group of elements the intensity or pitch is increased. Furthermore, the impression arises that the interval preceding the louder element is longer and the one following the louder element shorter. The lengthening of every second or third element results in the perception of "groups of two or three, with the shorter elements beginning the group" (Handel 1989:388). Changes in the frequency/pitch of an element can change the perception in several ways, e.g. the element with the highest pitch tends to be perceived as the accented one, and the least frequent element may seem to begin a group. Illustration of these aspects can be found in Handel (1989).

1.4.1 Stress–timing and syllable–timing

Isochrony has been referred to as "the intuition that stressed syllables come at equal intervals in time irrespective of the number of intervening unstressed syllables; interstress interval durations are equal" (Eriksson 1991:7). Accordingly Allen (1975, as cited by Zellner Keller & Keller, in press) notes a tendency to perceive period sequences much more regularly than they are in effect. Generally, long periods are underestimated and short periods are overestimated, resulting in a tendency toward "isochrony".

Even though the concept of isochrony has always been controversial, extensive research was carried out to find the acoustic correlates for the perceptual impression of isochrony (amonst others Uldall 1971, 1972, Lea 1974 (all as cited in Pompino-Marschall 1990), Dilley, Wallace & Heffner 2012).

Pike (1945) was the first to look for isochrony in different languages. He distinguished between stress- and syllable-timed languages, where for the stress-timed languages the inter-stress interval is the entity which shows isochrony, but for syllable-timed languages it is the syllable. Abercrombie (1967) agreed with Pike's theory and claimed that every language in the world belongs to one of these rhythm classes. According to Abercrombie, the difference between the languages is physiological, caused by the existence of two different muscular activities; i.e. responsible for the periodicity of syllables are the chest-pulses, for the periodicity of accents stress-pulses. However, electromyographical investigations of the respiratory muscles invalidated this claim (e.g. Ladefoged 1967, Lindblom 2010).

Nevertheless, the distinction between the rhythm classes had been widely accepted by researchers (amongst others Cruttenden 1997, Ball & Rahilly 1999, Rogers 2000 all as cited in Cauldwell 2002). It was assumed that there are distinct acoustic correlates to the perception of rhythm and research was carried out to classify the languages into the respective classes (see Dauer 1983 or Eriksson 1991 for a listing).

In this context, Abercrombie (1967) hypothesised that each language is assigned to one rhythm class, which shows either isochrony of stresses (in stress-timed languages) or isochrony of syllables (in syllable-timed languages). Additionally, the syllable length was said to vary in stress-timed languages, however, not in syllable-timed languages. In syllable-timed languages the interstress-interval length should be variable, but not in stress-timed languages (Abercrombie 1967, Cauldwell 2002).

However, a number of researchers (amongst others Roach 1982, Dauer 1983, 1987, Cauldwell 2002) argued against the hypotheses made by Abercrombie (1967). In order to elucidate that measurements of time intervals in speech could not provide evidence for the classification of languages into rhythm classes, Roach (1982) conducted an experimental investigation on six languages (three considered stress-timed (English, Russian, Arabic) and three syllable-timed (French, Telegu, Yoruba)). He investigated the assumption made by Abercrombie (1967) that stress-timed languages show great variability in syllable length and that syllables tend to be equal in syllable-timed languages. He found that results did not confirm Abercrombie's assumption. Roach (1982) declared that the criteria for the distinction in stress- and syllable-timed rhythm groups are deficient and may just be a matter of perception: "a language is syllable-timed if it sounds syllable-timed" (Roach 1982:78).

Another critic of the hypothesis made by Abercrombie was Dauer (1983, 1987). She carried out a comparative experiment of interstress interval duration in read speech in four different languages (English, Spanish, Italian, Greek). She did not find more regular interstress intervals in English (stress-timed language), than in Spanish (syllable-timed language).

Both authors (Dauer 1983 & Roach1982) state that there are no isochronous durations of interstress-intervals in stress-timed languages or of syllables in syllable-timed languages, and that there is very weak empirical evidence for isochrony. Therefore, the allocation to the different rhythm classes was based on intuition, as no empirical evidence could be found, e.g. "the regularity of the stresses is more apparent than real, in that listeners tend to perceive isochrony even in a sequences of inter-stress intervals that are manifestly far from equal" (Roach 1982:75).

Roach (1982) and Dauer (1983, 1987) therefore argued against a strict classification in stress- and syllable-timed rhythm classes. They suggest that languages could feature elements of both stress- and syllable-timing and that languages, depending on which of the rhythmic pattern is predominant, could be placed on a continuum where total stress- and syllable-timing represent the endpoints.

Roach (1982) elucidates that the allocation to different rhythm classes could not be made "on the basis of measurement of time intervals in speech" (Roach 1982:78), but that e.g. syllable structure, or vowel reduction needs to be taken into account. On the same note Dauer (1983, 1987) proposed that the impression of stress-timing and syllable-timing could be due to the combination of certain phonological, phonetic, lexical and syntactic properties in a language; in particular the presence or absence of vowel reduction, syllable-structure (e.g. complexity of consonantal clusters), and word stress. The syllable structure of stress-timed languages would be more variable than in syllable-timed languages, whereas vowel-reduction could be found in stress-timed languages but rarely in syllable-timed languages.

It is, therefore, assumed that there are possibly distinctive phonetic and phonological properties of syllable-timed and stress-timed languages. Ramus, Nespor & Mehler (1999) tried to connect the classification of the rhythm classes to acoustic parameters, with different elements recurring at regular intervals establishing temporal organisation: stresses in stressed-timed languages such as English, German or Swedish, syllables in syllable-timed languages such as French, Italian or Spanish.

Related to this topic Dauer (1983) had stated that stress-timed languages are said to show a larger variability of syllable duration and the stress is most often realised on the strong syllable. A strong syllable is a prominent syllable bearing the stress, which may be realized through changes in duration, loudness, and pitch (Dauer 1983). On the same note Clark & Yallop (1995) reported that in English, which is a stress-timed language, rhythm is established “by strong beats falling on the stressed syllables of the word” (Clark & Yallop 1995:340) which alternate with unstressed syllables. A typical spoken English utterance would consist of a number of rhythmic units, which each would be dominated by the beat of the stressed syllable. These units are organized to yield a regular rhythm, i.e. to keep the intervals between stress syllables isochronous. In the sentence CAR-la's GO-ing-to-the-pa-RADE (stressed syllables are indicated by bold letters) the unstressed syllables "ing", "to", "the", "pa" are likely to be articulated very fast to keep interstress interval durations equal and they may be reduced so that "is" or "to" can be pronounced without any vowel (Clark & Yallop 1995). In an utterance with more than one stress group, the stressed syllables would tend to occur at perceptually equal intervals in time.

Stressed syllables would show little or no reduction, whereas a tendency towards shorter or even absent unstressed vowels could be found in unstressed syllables to keep inter-stress interval durations equal (Clark & Yallop 1995). Therefore, stress-timed languages are said to have a greater degree of vowel reduction and a more varied syllable structure than syllable-timed languages (Grabe & Low 2002). Additionally, the "duration of inter-stress intervals in English" is reported to be "directly proportional to the number of syllables they contain" (Ramus et al. 1999:267). There are two causes that can influence the duration of an inter-stress interval, the position of the interval within an utterance as well as the type of syllable they contain (Ramus et al. 1999).

Different to stress-timed languages, syllable length is said to be more even and unstressed syllables are not expected to be reduced in languages with a syllable-timed rhythm (Kohler 1995, Wagner & Dellwo 2004). A smaller degree of vowel reduction or the absence of vowel reduction as compared to English is therefore said to be a distinctive feature of syllable-timed languages, e.g. French

(Fant, Kruckenberg & Nord 1991). Resilience against vowel reductions because of the stress/ unstressed contrast can be assumed.

As more and more languages and dialects are investigated perceptually and acoustically a clustering of different languages into syllable- and stress-timed groups is found and the categorical distinction into stress- and syllable-timed languages gets weaker (e.g. Ramus et al. 1999, Ramus 2002, Low et al. 2000, White & Mattys 2007a). Researchers assume that different languages can be more or less syllable- or stress-timed and possibly different degrees of syllable- and stress-timing occur (Roach 1982, Dauer 1983, 1987, Grabe & Low 2002, Ramus 2002, White & Mattys 2007a).

1.4.2 Rhythm Metrics

Recently, there has been a huge surge in studies that tried to find the acoustic correlates that form the basis of the perceived rhythmic differences and to find the best measure to capture the phenomenon. Studies focused on different elements in speech, such as syllables, inter-stress intervals, consonants, or vowels, to see which the element is that influences the perception of the different speech rhythms. In succession to that a number of different rhythm metrics have been proposed, where the authors tried to provide evidence for the traditional stress-timed and syllable-timed language dichotomy.

As detailed above, the rhythm of speakers with ataxic dysarthria differs from that of healthy control speakers, and researchers assume that it becomes more syllable-timed /less stress-timed as compared to healthy control speakers and cross-linguistic rhythm metrics can capture this difference (Liss et al. 2009).

The following section introduces different rhythm metrics[3]. The aforementioned details about ataxic dysarthria, as well as about rhythm, show that the perception of rhythm not only differs in different languages, but also in the speech of speakers with dysarthria and healthy controls. Some of the metrics introduced below were established for the investigation of dysarthric speech, others for cross-linguistic studies.

The measures are based on different measurement units to offer metrics of rhythm. They derive from the segmentation of speech into vocalic and consonantal intervals, syllables, or inter-stress intervals, used for measuring variability in these intervals. As detailed above, the hypothesis is that stress-timed languages show a tendency of greater contrast in vowel/syllable/inter-stress interval durations between stressed and unstressed syllables, whereas syllable-timed languages would show little variability in succeeding vowel/syllable/inter-stress interval durations.

Current research developed and used several rhythm metrics based on vocalic and intervocalic interval durations. The following table 3.1 provides a summary of the main metrics (see also Wiget et al. 2010, Knight 2011, Lowit 2014):

[3] Only rhythm metrics important for the present study will be introduced in detail. It would be outwith the scope of this Master thesis to include all available rhythm metrics.

Table 1.3 Overview of prominent rhythm metrics

	Rhythm Metric		Main reference
Vowel based	%V	See below table 1.4.	Ramus et al. 1999
	ΔV	Standard deviation of vocalic interval duration.	Ramus et al. 1999
	VarcoV	See below table 1.4.	White & Mattys 2007
	PVI	Normalized pairwise variability index: Mean of the differences between successive vocalic intervals divided by their sum, mulitplied by 100	Low et al. 2000
	nPVI-V	See below table 1.4.	Grabe & Low 2002
Consonantal based	ΔV	Standard deviation of vocalic interval duration.	Ramus et al. 1999
	VarcoC	Coefficient of variation of consonantal interval duration multiplied by 100.	White & Mattys 2007
	rPVI-C	Pairwise variability index for consonantal intervals. Mean of the differences between successive consonantal intervals.	Grabe & Low 2002
Vocalic & consonantal based	nPVI-VC	Normalized pairwise variability index for summed vocalic and consonantal intervals. Mean of the differences between successive vocalic+consonantal intervals divided by their sum, multiplied by 100.	Liss et al. 2009
Syllable based	VarcoVC	the normalized standard deviation of the duration of successive combined vocalic and consonantal intervals, as an approximation to syllable duration	Liss et al. 2009 Deterding 2001
	VI	See below table 1.4.	Deterding 1994
Stress group based	ISI	See below table 1.4.	Fant et al. 1991

The main metrics presented here and used in the current study are: The proportion of vocalic intervals (%V) (Ramus et al. 1999), normalized Pairwise Variability Index (nPVI) (Low et al. 2000), VarcoV (White & Mattys 2007a), Variability Index (VI) (Deterding 2001), the coefficient of variation of inter-stress intervals (ISI) (Fant et al. 1991). The first three metrics are based on vowel durations (%V, nPVI, VarcoV), whereas the VI is a syllable based measure. The ISI is based on inter-stress interval durations.

Table 1.4 Rhythm metrics of the study

Rhythm Metric		Main references
nPVI	$$nPVI = \sum_{i=1}^{n-1} \left[\frac{\frac{\| x_{(i)} - x_{(i+1)} \|}{\left(x_{(i)} + x_{(i+1)}\right)/2}}{n-1} \right] x100$$ Normalized pairwise variability index: Mean of the differences between successive vocalic intervals divided by their sum, mulitplied by 100.	Low (1998), Low et al. (2000)
%V	Percentage of vocalic intervals: Percent of total utterance duration composed of vocalic intervals.	Ramus et al. (1999)
VarcoV	Coefficient of variation of vocalic interval duration: standard deviation of the vocalic interval divided by the mean vocalic interval duration and multiplied by 100	Ramus et al. (1999), Dellwo (2004), White & Mattys (2007a)
VI	Variability Index: Normalisation: $z_{(i)} = \frac{x_{(i)}}{\bar{x}_{(n-1)}}$ Rhythm measure: $$VI = \sum_{i=1}^{n-2} \frac{\| z_{(i)} - z_{(i+1)} \|}{n-2}$$	Deterding (1994), Low (1998)
ISI	Coefficient of variation (CV%) of duration of inter-stress intervals	Fant et al. (1991), Hartelius et al. (2000)

In the following sections the rhythm metrics will be introduced in more detail.

1.4.2.1 %V

Ramus et al. (1999) first proposed the %V, ΔC and ΔV for the investigation of cross-linguistic differences. They developed a classification of rhythm classes based on the instrumental measurements of vowel and consonant durations.

They examined 160 utterances of four speakers of eight different languages (English, Dutch, Polish, French, Spanish, Italian, Catalan and Japanese). They measured vocalic and consonantal intervals within an utterance. A vocalic interval was between the beginning (onset) and ending (offset) of a vowel (Ramus et al. 1999). A consonantal interval was set between consonantal onset and offset or cluster of consonants. No individual phoneme durations were measured, but only the duration of the vocalic or consonantal intervals. The total duration of an utterance could therefore be added up by the sum of the duration of all vocalic and consonantal intervals in a sentence.

The authors tried to provide evidence for the traditional stress-timed and syllable-timed language dichotomy and claim that ΔC (standard deviation of the consonantal intervals), ΔV (standard deviation of the vocalic intervals) as well as %V (proportion of vocalic intervals) are directly related to syllabic structure of languages. Results showed %V to be the best acoustic correlate of rhythm class. Stress-timed rhythm showed the higher %V, and syllable-timed languages a smaller %V value. The authors argue that %V and ΔC were directly related to structure of a syllable, and more syllable variety would lead to heavier syllables. According to Ramus et al. (1999) a heavy syllable would include more consonants, therefore the duration of the consonants within a syllable would be higher, and the duration of the vowel therefore shorter. Hence, a higher ΔC and smaller %V.

Some studies doubt the reliability of Ramus et al.'s (1999) measures and their ability to indicate difference in rhythm classes and additionally criticized it as not being stable against influences of articulation rate and not describing rhythm but rather syllable complexity (Low et al. 2000, Dellwo 2004 ,Wagner & Dellwo 2004). Wagner & Dellwo (2004) found the measures as not revealing information about rhythm as being a "regularly ordered sequence of events in time" (Wagner & Dellwo 2004:227). The three metrics (ΔC, ΔV, %V) do not account for the sequence of intervals and no information is given about the durational difference of the succeeding intervals, therefore they are criticized as only measuring the total time of vocalic intervals within an utterance, compared to consonantal intervals. The %V measure, however, could probably be an indicator of syllable complexity, based on the assumption that greater syllable complexity results in proportionally more vowels and fewer consonants per syllable. Therefore, some studies found that %V was the only one of the three metrics to be a good indicator of rhythm class (Ramus et al. 1999, Ramus 2002, Dellwo & Wagner 2003, White & Mattys 2007a, Mok & Dellwo 2008, Liss et al. 2009), assuming that in stress-timed languages a higher degree of vowel reduction takes place and syllable complexity is higher.

The %V does, in contrast to other rhythm metrics, not include a normalisation of articulation rate. Normalisations are included in rhythm metrics to compensate for possible influences of articulation rate. Vowels and consonants are often reduced or omitted at higher speech rates. Furthermore, rhythm metrics that are not controlled for speech rate, could reflect the individual speaker idiosyncracy, as well as languages typologies (White & Mattys 2007a). As speakers with ataxic dysarthria often show significantly slower articulation rates compared to healthy control speakers (Kent et al. 1997, Cannito & Marquardt 1997, Duffy 2012, Brown et al.1970), it is important that rhythm metrics applied to ataxic dysarthric speech are stable against articulation rate.

White & Mattys (2007a) did not find a correlation with articulation rate for the %V metric. Dellwo & Wagner (2003) only found little correlation in their 2003 study, but a high correlation in their 2004 study (Wagner & Dellwo 2004) and suggested that data investigated with the %V needed to be controlled for speech rate.

Knight (2011) tested the reliability and validity of seven different rhythm metrics (nPVI-V, rPVI-C, ΔC, ΔV, %V, VarcoV and VarcoC) across time for the same speaker and the same task. Results showed that %V was the measure the most valid and reliable across time compared to the other measures.

Liss et al. (2009) applied amongst others the %V to the speech of ataxic dysarthric speakers and found it suitable to distinguish between their dysarthric group and control group. However, they also found high correlations between %V and speech rate and therefore did not find it as most useful metric for the differentiation of different dysarthria groups.

Arvaniti (2012) tested six different rhythm metrics (ΔC, %V, rPVI, nPVI, VarcoC and VarcoV) on spontaneous speech, sentences and story reading on six different languages (English, German, Greek, Italian, Korean and Spanish) with eight speakers per language. Results did not show the typical stress- and syllable-timed classification of languages. She found inconsistent results across all tested metrics and mostly no significant task differences. Furthermore, she detected high inter-speaker variation and task dependency.

1.4.2.2 Normalised Pairwise variability Index - nPVI

The use of a rhythm index was first suggested by Francis Nolan, which was then established in an MPhil thesis by Low in 1994. The rhythm metric was thereupon further developed by Low & Grabe (1995), and Low (1998), who extended their previous work by including a pairwise normalisation. The current study uses the nPVI as developed by Low (1998) and published by Low et al. (2000). Low et al. (2000) wanted to test if successive vowel durations are more equal in Singapore English (syllable-timed language) compared to British English (stress-timed language). They tested their metric on two sets of sentences with five speakers of Singapore English and ten speakers of British English. The normalised Pairwise Variability Index (nPVI) mirrors changes in successive vowel length over an utterance independent of the speaking rate and is derived from successive vowel durations. It focuses on the sequential nature of rhythm. The normalisation procedure integrated within the nPVI is derived by dividing the difference between the pairs of vocalic intervals by the sum of the intervals.[4] This has been found useful by some researchers, since no correlation of the nPVI values and articulation rate were found (White & Mattys, 2007 a, b, Liss et al. 2009).

The nPVI has been used for the investigation of cross-linguistic rhythm classes and showed that it is a reliable index of rhythmic differences across languages (Low et al. 2000, Spencelayh 2001, White & Mattys 2007a, Mok &

[4] For a detailed description of the formula and application procedure of the nPVI to speech data see chapter two methodology.

Dellwo 2008, Liss et al. 2009), only Arvaniti (2009, 2012) did not find expected statistically differences.

Different studies showed that the nPVI is not only able to distinguish between different rhythm classes, but is also sensible for the differentiation within the rhythm classes able to distinguish different dialects of English (e.g. Low et al. 2000, Spencelayh 2001, White & Mattys 2007a). This shows that a gradient and not categorical distinction between the rhythm classes may be possible. It also indicates that the measure might be able to capture differences between ataxic dysarthric and healthy control speech, as it could be argued that if a measure captures dialectal differences in one language, it may also be able to capture the difference between healthy and disordered speech. Knight & Cocks (2007) tested the nPVI on the speech of a speaker with right hemisphere damage and a control speaker and found a significant difference between the speakers.

The capability to differentiate healthy from dysarthric speech was shown in other studies: The nPVI has been successfully used for the differentiation of normal from ataxic speech (Liss et al. 2009, Stuntebeck 2002, Rosen, Kent & Duffy 2003) and it has been shown to be a useful tool in "capturing the rhythmic properties of ataxic speech" (Stuntebeck 2002:22). Furthermore, the nPVI was also able to distinguish between different types of dysarthria. However, Liss et al. (2009) did not evaluate the nPVI to be the best metric for the differentiation between dysarthria types because it showed less classification accuracy between dysarthric speaker groups than other metrics.

1.4.2.3 Variability Index – VI

Deterding (1994) developed the Variability Index (VI) first introduced by Low (1998) by applying "a normalization procedure based on the syllable rate of the utterance" (Deterding 2001:218-219). The VI is based on differences in the duration of neighbouring syllables, with the exclusion of the final syllable and normalization based on the whole utterance. The VI is calculated by "subtracting the duration of one syllable from that of the next syllable" (Deterding 2001:222) and the application of the normalization, which "was achieved by dividing the duration of each syllable by the average duration of all the syllables (except the last) of the utterance" (Deterding 2001:222). The authors argue, that vowel-less syllables are quite common in English, i.e. the word "perception" is perceived as a three syllable word, even though in the final syllable there may be no vowel (/pər'sepʃən/ versus / pə(r)'sepʃn/). The [n] would become syllabic. Even though, it may be argued, that vowel-less syllables are not too common, instead vowels are reduced, the authors reason that it is more feasible to use a measurement which is based on the whole syllable. Results of the study of Deterding (2001) with six speakers of Singapore and British English each showed that the VI was able to distinguish between Singapore and British English and found the VI to be a suitable measure for the differentiation in rhythm along a stress- and syllable-timed scale and this on conversational data.

1.4.2.4 VarcoV

The VarcoV emerged from the ΔC and ΔV measures of Ramus et al. (1999) which were further edited by Dellwo (2004) and finally by White & Mattys (2007a). Even though the %V and standard deviations of consonantal intervals ΔC were found to be two dimensions to differentiate rhythm classes on the acoustic level by (Ramus et al. 1999), Dellwo (2004) argued that especially the ΔC was influenced by speech rate (see also Barry, Andreeva, Russo, Dimitrova, Kostadinova 2003 and Dellwo & Wagner 2003). He therefore established the VarcoC which is a measure of relative variation, "calculated as the percentage of the standard variation of the consonantal interval duration (ΔC) of the average duration of consonantal intervals (meanC)" (Dellwo 2006:232-233). He claims that rhythm classes would be better differentiated with this normalised measure VarcoC, since some languages vary in rhythm with regard to speech rate while others seem to stay unaffected.

White & Mattys (2007a) then used the rate normalised variation coefficient of the duration of vocalic intervals (VarcoV) and found it to be discriminative for rhythmically distinct languages and robust to articulation rate variation (White & Mattys, 2007a). They used speech from six speakers of English, Dutch, Spanish and French, who read five sentences in their own language and one of the other languages they were competent in. To test for rate dependency they calculated correlations between the VarcoV and articulation rate.

Whereas some studies report the VarcoV to be robust against articulation rate, Liss et al. (2009) found a high correlation with articulation rate and excluded the VarcoV from their analysis of dysarthric speech, even though it can successfully be used for the differentiation between ataxic and normal speech.

Further studies confirmed the usability of the VarcoV for the distinction between stress- and syllable-timed languages (Dellwo 2008a,b, White & Mattys 2007 b), as well as distinguishing between native and non-native speakers of English (White & Mattys 2007a,b). Only Arvaniti (2009, 2012) could find no significant effects of language in their investigation of six different languages.

White & Mattys (2007 a,b) used several different rhythm metrics and also proposed the VarcoV as a normalised variation coefficient of the duration of vocalic intervals (see 1.4.2.5). For the current study the VarcoV was chosen, as it was able to distinguish between rhythm classes in the studies by White & Mattys (2007 a,b) and they found it to be stable against articulation rate.

1.4.2.5 Coefficient of variation of inter-stress intervals - ISI

The basis of the inter-stress intervals are feet, i.e. one stress interval contains one foot, which consists of one stressed syllable and a number of unstressed ones until the next stressed syllable in the speech segment. The intervals between main stresses (Inter-stress Intervals) are measured in the study of Hartelius et al. (2000). The coefficient of variation across the Inter-stress Intervals (ISI) can indicate a tendency towards isochrony of stress groups.

Fant et al. (1991) applied the ISI on English, Swedish and French speech data and found it harder to identify major stresses in French, compared to English and Swedish, which is plausible because French is a syllable-timed language.

They concluded that the investigation of inter-stress intervals may not be sufficient as a basis for describing different languages.

Hartelius et al. (2000) investigated a reading passage and sentences of fourteen Swedish patients with ataxic dysarthria. The authors found significantly longer syllable duration, significantly longer mean durations of inter-stress intervals in the ataxic dysarthric group compared to the control group. The variability of the duration of the syllables showed higher intra-utterance variability for the control group, but a significantly higher inter-utterance variability for the syllable duration and a significantly higher variability of the duration of inter-stress intervals for the ataxic dysarthric group. This indicates a notably higher variability for their Swedish speakers with ataxic dysarthria than for their controls (Schalling & Hartelius 2004). Schalling & Hartelius (2004) furthermore found a correlation between the severity of the neurological impairment and the performance in the speech tasks. Hartelius et al. (2000) conclude that the absence of even inter-stress intervals (regular stress beats) adds to the perceptual impression of scanning speech.

The ISI has been demonstrated to be able to indicate a tendency towards increased variability of ISI in patients with spinocerebellar ataxia (Hartelius et al. 2000, Schalling & Hartelius 2004) in Swedish. Swedish is characterised by prominent pitch accents and sentence stress is more predictable than in English. There is thus a question whether this measure can be applied to English speaking participants, particularly in more naturalistic speech tasks where stress placement is less predictable.

As illustrated above a range of different rhythm metrics have been established for the investigation of different languages, as well as dysarthric speech. Particularly the %V, nPVI, VarcoV, VI and ISI have been introduced, as they are based on different elements (vowels, syllables, interstress- intervals) and have shown to distinguish different speakers. However, it has not been tested for all of them, if they actually measure what is heard as a different rhythm. As it is important that a metric actually measures what is heard, the following section will give information on perceptual evaluation of dysarthric speech.

1.5 Perceptual Evaluation of dysarthric speech

Perceptual assessments have a long tradition in dysarthric research (Darley et al. 1969a, b). Scherer & Ekman (1982) state that "almost all of the objective parameters of an acoustic speech wave form can be "heard" by judges and can consequently be assessed with the help of category systems and rating scales" (Scherer & Ekman 1982:169).

However, with the increasing availability of instrumental methods some researchers have challenged the value of perceptual analysis. The two main points of criticism are the lack of reliability in ratings (inter- and intra-rater reliability) as well as limited analytic potential (Hartelius, Theodoros, Cahill & Lillvik 2003). Results of perceptual analyses may not highly correlate with the output of acoustic measures and acoustic dimensions can often be defined more precisely than perceptual impressions of speech data. The perceptual impressions may furthermore often refer to more than one acoustic correlate (Patterson 2000, Lowit-Leuschel 1997).

One strong argument to perform perceptual analysis is that perception experiments capture what listeners actually hear. Even though the agreement of listeners' perception differs with regard to which speech dimension is being rated, the agreement is quite high for overall speech dimension (such as intelligibility or naturalness of speech) (Linebaugh & Wolfe 1984, Hartelius et al. 2003). The degree of agreement lowers the more specific the rated dimension is (i.e. irregular articulator breakdown, distorted vowels etc) (Linebaugh & Wolfe 1984, Hartelius et al. 2003). The importance of the perceptual impression of a person's speech is such that it is the impression which "ultimately defines if, in which parameters and to what degree speech is impaired" (Lowit-Leuschel 1997:36).

Two different methods can be used for the perceptual evaluation of speech: categorical rating or the magnitude rating. For the magnitude rating listener judges have to rate speech characteristics as an estimate, with no preset categories given. Instructions are given to the listener judges, which explain that the response should be proportional to the intensity of the experienced speech stimulus (Patterson 2000[5]). Often preference is given to the magnitude rating because it is said to be a more direct measurement of a subjective experience or sensation. On the other hand, it is criticised because a linear relationship between the rating and the magnitude of the subjectively experienced sensation

[5] The following paragraph refers mostly to Patterson 2000, if not stated otherwise.

may not be given or may be distorted (Patterson 2000). Though researchers have used the direct magnitude estimation successfully (Setter, Stojanovik, van Ewijk & Moreland 2007), a direct relationship between rating and the subjectively experienced sensation is not given by magnitude rating, but by the category rating.

Category ratings come in many different forms: similarity judgements, typicality judgements, psychophysical scales, health self-reports, attitude questionaires inter alia (Petrov & Anderson 2005). These are mostly used for measuring direction or strength of subjective states, represented by a rating scale where each end shows the end of a continuum (weak, strong, not at all confident, and confident). Hereby an "ordered set of categories" (Petrov & Anderson 2005:383) is used to rate cases. A well known category rating scale is the Likert scale, a prototype of row of boxes with labels defining each end of the continuum (Patterson 2000:52).

Whilst it is stated as a critique that category ratings may produce estimate ratings of the discriminability of stimuli rather than "estimates of perceptual magnitudes of the stimuli used" (Patterson 2000:52), its simple instructions are usually easy to understand for the listener judges. In this context, another factor which needs to be accounted for is the ability of participants to carry out ratings. "Unwillingness to expend cognitive effort" (Patterson 2000:56) may have an impact on the results of ratings. Therefore, the less complex instructions of category ratings as opposed to magnitude estimates may be an advantage. It is further criticised that the measurements of subjective states collected through a category rating are not absolute, however, this also holds true for magnitude rating; and therefore the one cannot be favoured over the other.

In addition to the fact that category ratings are easy to make, conduct and understand and even large amounts of data can be efficiently collected and analysed, there is a long tradition of the use of category ratings in speech research.

1.6 Studies in atypical population

As previously detailed there is a wide variety of rhythm measures available that have been developed for different speaker groups and speech materials and are based on different algorithms. It is difficult to know how much the results of previously published studies are comparable and can be generalised. Despite the fact that individual measures have been criticized by authors who presented alternatives to that measure, only very few studies have yet performed a comparison of a wide range of measures across a variety of speech tasks and languages to provide information on this issue.

Liss et al. (2009), however, carried out a study with 55 speakers with different dysarthrias: twelve with ataxic dysarthria, nine with hypokinetic dysarthria, twelve with hyperkinetic dysarthria, ten with a mixed spastic-flaccid dysarthria and twelve healthy control speakers. Their aim was to test whether control and dysarthric speech and different dysarthria types could be discriminated by durational measure of rhythm. They collected vocalic and consonantal segment durations and applied ten different rhythm metrics (ΔV, ΔC, %V, VarcoV, VarcoC, VarcoVC, nPVI-V, rPVI-C, nPVI-VC, rPVI-VC (see table 1.3) to it, and also measured articulation rate. Liss et al. (2009) chose five sentences which were previously used by White & Mattys (2007a, b) and therefore allowed a direct comparison of the results. Additionally, they collected a set of 80 phrases which were developed on a larger study in dysarthric speech (Liss, Spitzer, Caviness, Adler, & Edwards 2000). First analysis showed that all metrics were able to show group differences. Additionally, through a stepwise discriminant analysis they tested which metric would best classify the different dysarthric speech groups. Furthermore, a non-stepwise discriminant analysis was carried out on the 80 phrases as a cross-validation. For the various combinations of the different rhythm metrics used, their results showed a high over all classification accuracy into speakers groups (79-88%). Here, Liss et al. (2009) found VarcoV, VarcoVC, ΔV, %V and ΔC as the best set of rhythm metrics for distinguishing the speakers into the groups. The further analysis was carried out on dysarthria specific comparisons. These, however, showed different metrics to best discriminate specific types of dysarthria, e.g. VarcoV for mixed flaccid-spastic dysarthria, or the VarcoC, rPVI-VC and nPVI-V for ataxic dysarthria.

Taken the results of the Liss et al. (2009) study, a rhythm metric that is "sensitive to the particular pattern of rhythm generated" (Liss et al. 2009:1345) by the speaker group should be used for any new research. For the investigation of ataxic dysarthric speech, the VarcoC, rPVI-VC and nPVI-V (here nPVI) would be suitable metrics according to Liss et al. (2009).

Other studies also found the nPVI suitable for the investigation of ataxic dysarthric speech. Therefore, a study of ataxic dysarthric speech could also include metrics that have shown to be sensitive rhythmic differences in dysarthric research, such as nPVI, %V, VarcoV on the one hand, but also e.g. syllable or interstress intervall based metrics such as VI or ISI. ISI has successfully differentiated ataxic from control speech in previous studies (White 2012, Hartelius et al. 2000, Schalling & Hartelius 2004). Liss et al. (2009) preferred rhythm metrics based on vocalic and consonantal intervals and not syllables, as they state that the rhythmic abnormalities in dysarthric speech are at a level of articulatory implementation and are not arising from phonological constraints.

However, syllable-based or interstress interval based metrics may reflect the perceptual impression of deviant rhythm in ataxic dysarthric speech, as some intervals or syllables stand out more perceptually than others.

Whilst the studies above have contributed some important new knowledge, they can be critised on a number of aspects. For example, not many studies compared performance across different speaking tasks. Liss et al. (2009), Low et al. (2000) and Lowit (2014) only applied the rhythm metrics to two different speech tasks. This allows the investigation of larger corpora and read material allows for a good comparability. However, the use of a wider range of different speech tasks could show possible task dependent performance which is a frequent symptom of disordered speech (see above). This is particularly the case for rhythm as different speech tasks can show considerable variations in rhythmic structure even within the same language, as demonstrated by Low et al.'s (2000) stressed and unstressed sentence sets.

A second limitation of many of the above studies is the lack of perceptual validation of the acoustic metric results. Liss et al. (2009), for example, investigated a significant number of rhythm metrics to find out which could differentiate speakers with different types of dysarthria. However, they did not include perceptual analysis in their study. The correlation of the results of a perceptual evaluation of the speakers' rhythm and the results of the rhythm metrics can show if the metrics actually measures what is heard. Only very few studies carried out such a correlation (Stuntebeck 2002, Lowit 2014) and showed differing results. Whereas Stuntebeck (2002) found a connection between the perceived rhythmic disturbance and nPVI results, Lowit (2014) did not. Her results for the nPVI, as well as VarcoV and %V (amongst others) showed only a poor relationship between the results of the perceptual and rhythmic measures. If the application of rhythm metrics is to help find out more about the rhythmic performance of disordered speakers and possibly have some clinicial application at some stage, it is essential to validate the acoustic rhythm metrics perceptually to ensure that they really mirror what is perceived as a deviant speech rhythm rather than some other aspect of speech disturbance (see also Lowit 2014).

1.7 Research Questions

This study investigates the differences in speech rhythm between healthy control and ataxic dysarthric speakers of American English across a variety of different speech tasks and rhythm metrics with the aim to further validate Liss et al.'s (2009) findings on which acoustic rhythm metric is most suited to differentiate these two speaker groups, but in addition to address some of their shortcomings by also focussing on which task(s) can highlight such differences most, and which rhythm metrics correlates best with perceptual evaluations.

The following research questions were asked to pursue these aims:

1. Which rhythm metric is most suited to highlight rhythmic disturbances in ataxic speech compared to healthy control speech?
2. Which speech tasks are most sensitive to capture rhythmic changes in ataxic dysarthric speech?
3. Which rhythm metric best reflects the perceived impression of the speech disturbance of the ataxic dysarthric speakers?

The current study is highly explorative, looking at a wide range of measures and tasks, and correlating results with a perceptual evaluation. As the existing literature is relatively contradictory, no hypotheses can be provided.

The next chapter provides a detailed description of the methodology used in this study to address these research questions.

2. Methodology

This chapter reports on the methodology used for the investigation of the research questions of this study. The study focused on the testing of suitable rhythm metrics to highlight rhythmic disturbances in ataxic dysarthria across a range of speech materials, which was also correlated with the perceptual evaluation of the participants' speech rhythm. A summary of the speakers' details are given, as well as the selection criteria which were applied. The materials and the rhythm metrics are described and the reason for their inclusion is given. The recording procedure is reported, followed by a depiction of the perceptual as well as the acoustic analytic procedures applied to the data.

2.1 Participants

Six American English (AE) speakers (living in the US) with a confirmed medical diagnosis of spinocerebellar or cerebellar ataxia and dysarthria took part in the study, along with six control participants who were individually matched for language, gender and age. For reasons of confidentiality and for reference all speakers were given labels indicating their language background and whether they were ataxic dysarthric or control speakers, and a number for identification.

Example: <GroupLanguage_Number>

Group: A= speakers with ataxic dysarthria, C = control speakers

Language: E = English

AE_01 – AE_06, CE_01 – CE_06

All participants were informed about the procedures of the study. In line with Strathclyde University and NHS ethics guidelines, written consent was gained for their speech to be recorded, analysed and results to be published in anonymised way.

The sample size of the current study is relatively small, which was necessary to remain within the scope of the thesis but maximise the number of tasks and measures required for the intended comparison. Although the group sizes are marginally smaller that in the Liss et al. (2009) study, it still compares favourable with many other studies published on rhythm both in the disordered and cross-linguistic field.

2.1.1 Selection criteria of the participants

In order to keep the ataxic dysarthric sample as homogeneous as possible only speakers with cerebellar ataxia were selected. The speakers with ataxic dysarthria (AE_01 to AE_06) were all patients at the Neurologic Clinic at the University of California in San Diego, USA (see table 2.1). They were recruited for another study on speech in spinocerebellar ataxia (SCA) through the neurologist in charge of their care. The data was collected by an experienced speech and language therapist. The ataxia was diagnosed by neurologists, and the SCA was tested and specified mostly with the commercially available genetic tests for SCA (see chapter 1 on degenerative ataxia).

Results showed that only speaker AE_01 showed negative results for all available genetic tests. However, she was given the diagnosis of cerebellar ataxia/cerebellar degenerative disease based on the clinical picture and the MRI results (hypoplasia of cerebellum and brainstem) and could be included in the study.

Furthermore, all ataxic speakers had to be diagnosed as showing signs of mild to moderate dysarthria. An experienced speech and language therapist at the clinic carried out a perceptual identification of the presence and type of the speakers' dysarthria. The speech and language therapist looked especially for the presence of excess and equal stress, and scanning speech.

An exclusion criterion for the ataxic dysarthric participants was if the dysarthria was accompanied with other speech and language problems, such as apraxia of speech or aphasia. Furthermore signs of cognitive problems or depression would have led to an exclusion of the participant, since they can have an effect on speech prosody (Mundt, Snyder, Cannizzaro, Chappie, & Geralts 2007). Information of the treating neurologist and speech and language therapist was taken as a basis for the presence of these exclusion criteria.

All participants were informally tested by the speech and language therapist to have adequate reading skills and no severe hearing or visual problems which would keep them from reading and understanding the materials.

Participants with ataxic dysarthria as well as control speakers were recruited in a wide age span and from both genders. Since ataxic dysarthria is present in young as well as elderly speakers, no age limit was set for the participants of the study given the rarity of the disorder (incidence of the SCAs is estimated 5 in

100.000 persons (Evidente et al. 2000), the prevalence is reported as 3 in 100.000 (Schöls et al. 2004)) and the resulting difficulty in recruiting suitable participants. The wide age range was not expected to impact on the results as studies have reported differences for old and young or male and female speakers for voice quality and rate, but not for rhythm (e.g. Amerman & Parnell 1992).

2.1.2 Detailed participant information

The group of speakers with ataxic dysarthria consisted of two male and four female speakers, age range from 27 to 41, mean age 32.5 (table 2.1). The participants' dysarthria was mostly mild and intelligibility levels ranged from good (96% for AE_02 and AE_03) to poor (66% for AE_01).

Intelligibility levels were acquired by means of perceptual evaluation of the participants' speech. Listener judges assessed the speakers with respect to their intelligibility on the Poem with the scaling identification method (for a detailed discussion see Kent, Weismer, Kent & Rosenbek 1989) (see 2.4.1 for details). Here, listeners have to select a rating within a continuum of intelligibility. For example, the speech samples can be evaluated with descriptors, such as "always", "sometimes", "seldom", "not at all" to describe intelligibility or using a numeric scale, such as 1-100, where 100% represented readily intelligible and 0% not intelligible (unintelligible) at all. An advantage of the scaling method is that it can relatively easily be conducted. Whilst the subjective nature of the ratings (i.e. different raters may have different internal rating criteria) is often mentioned as disadvantage, the scaling method can be conducted relatively easy.

The intelligibility of the speakers was rated using a numeric scale in percent, where 100% represented highly intelligible and 0% not intelligible (unintelligible) at all. For this each speech file was played to the students once. If any of the students indicated that they needed to hear it again, it was played again.

Table 2.1 Details of the speakers

Speaker	Gender	Age	Diagnosis	Intelligibility in %
AE_01	F	28	Cerebellar Ataxia	66
AE_02	F	41	SCA 3	96
AE_03	M	29	SCA 3	96
AE_04	M	39	SCA 3	80
AE_05	F	27	SCA 8	84
AE_06	F	31	SCA 8	91
Ataxic dysarthric	Age range	27-41	Mean: SD	32.5 5.4
Control	Age range	23-39	Mean: SD	32.5 5.4
Abbreviations: AE= English speaking ataxic dysarthric speaker, SD= standard deviation				

Control participants were selected to match the ataxic dysarthric speakers in terms of age, gender, and variety of English (American English). The group therefore consisted of two male and four female speakers, age range from 23 to 39, mean age 32.5. Exclusion criteria for the control participants were any history of neurological impairment, and a history of or current speech and language problems.

The control participants were not recruited in the USA. Nevertheless, all control speakers were native speakers of American English, who were raised in the US and had lived there for most of their lives. At the time of their recording they were based in Glasgow, Scotland due to study or work commitments. However, they had lived in Glasgow for no longer than 1.5 years. Whilst Smith & Rafiqzad (1979) posit that living in a linguistic community for more than four consecutive months may have an effect on the linguistic behaviour of speakers, the current study nevertheless decided to use these participants because a strong American accent was perceived for all speakers by the examiner and experienced listeners. In addition, it was assumed that the changes in their linguistic behaviour may not have been as significant as in the Smith & Rafiqzad study, which investigated non-native speakers of English and thus had more scope for change than might be assumed between two varieties of English. This assumption was further supported by the fact that the speakers were all active in the American community in Glasgow, i.e. most of their contacts were Americans. Finally, the author checked the American English accent by comparing the perceptual impression of rhythm of the speaker group with another Scottish English group for similar tasks, i.e. Spontaneous Speech. Statistics showed marginal significant different between the groups (z=-2,025, p= 0,043*), indicating that the American English speakers did sound different, however close to the Scottish English speakers, given that they were both speaking English. Additionally, the American English speaker group and Scottish English group did not differ with regard to rhyhtm metric resuls (VarcoV & nPVI e.g. for the Spontaneous Speech sample: VarcoV z=-0,129, p=0,897, nPVI z=-0,129, p=0,897) and they fell in the range of previously reported results (Liss et al. 2009).

2.2 Materials

The following section reports on the material used in this study. Seven different speech tasks were chosen (see table 2.2). They range from highly structured to unstructured tasks. Advantages of structured tasks are usually the high degree of comparability of the speech output of different speakers. Less structured tasks may on the other hand be closer to the spontaneous speech behaviour of the participants.

The structured tasks were designed to capture a range of increasing predictability of stress placement and included spontaneous speech, a reading passage, sentences as well as a poem with a very specific stress pattern. A full list of all tasks for the study is provided in table 2.2.

Table 2.2 Material of the study

Material	
Poem	Reading of a Limerick from Robertson & Thomson (1986) at habitual pace
Sentence Reading	Seven sentences read in fast (maximum) and habitual pace from Lowit et al. (2001)
Full and Reduced Vowel Sentences	Reading of a set of five potentially full and five potentially reduced vowel sentences from Low et al. (2000) at habitual pace
Reading Passage	Reading of a dialogue-like Reading Passage adopted from Lowit-Leuschel & Docherty (2001) at habitual pace
Spontaneous Speech	Sample from a naturalistic speech sample, participants told about a restaurant visit

The same tasks and procedures were applied for the ataxic dysarthric and the control group. The read material allowed for grammatical and lexical control and, as mentioned above, good comparison of the data. The spontaneous speech task was included to elicit a speech sample as naturalistic and close to normal every day speech as possible.

All material was presented to the speaker in written format in a folder on a table in front of them. Speakers were asked to make sure they were comfortable before the recording started.

The following sections give detailed information about each task. Details about measurements applied to the tasks and statistics are given later on in this chapter (section 2.4.3).

2.2.1 Poem

The Poem taken for this study was a Limerick (Robertson & Thomson, 1986, see table 2.3). The participants were asked to read the poem at their habitual speed and repeated the whole line or whole poem if they made a mistake.

Table 2.3 Poem (Limerick)

"There was an old man with a beard
Who said, It´s just as I feared!
Two owls and a hen,
Four larks and a wren,
Have all built their nest in my beard!"

The Poem is made up of five lines with set metrical feet in each: three in the first, second and fifth and two metrical feet in the third and fourth lines. It includes 35 vowels, twelve Inter-Stress-Intervals, and 35 syllables for analysis. The number of syllables could change due to two reasons: Some speakers did not produce each word or syllable due to their speech problem or natural connected speech processe. In addition, the word "feared" and "owls" were produced with as a two syllable word by some speakers (e.g. AE_04, AE_06). Other speakers read "it was", "it is" or "I was" instead of "it's" (AE-02. AE_04, AE_05, AE_06).

Research reported about identified performance differences between highly structured and more natural materials like reading or conversation (Ziegler & Wessel, 1996, Ziegler 2002, 2003). The Poem is a highly structured task in that it is eliciting the strictest rhythm. As detailed in chapter 1 English is a stress-timed language where a reduced vocalic system in unstressed syllables and a tendency towards shorter or even absent unstressed vowels can be found and therefore a great variability of syllable duration between stressed and unstressed syllables (Dauer 1983, Grabe & Low 2002). Since stress placement is mostly isochronus in the Poem and not in more spontaneous speech. The difference between stressed and unstressed syllables can be expected to be bigger in the Poem than for other tasks, due to vowel reduction on unstressed syllables.

The difference between the group of speakers with ataxic dysarthria and the control is expected to be significant, as Brown & Docherty (1995) report an increase of unstressed vowel duration in read speech for speakers with dysarthria.

2.2.2 Fast and habitual pace sentences

The sentences used for the fast and habitual pace reading were taken from another study on VOT. They were controlled for a variety of initial plosives and include all English vowels (Lowit, Miller, Poedjianto, & McCall 2001). The initial plosives allow for a good acoustic analysis and the use of different vowels makes for a balanced speech task. They included mostly monosyllabic words and all sentences altogether had 65 syllables (see table 2.4).

Participants were asked to read the sentences once at their habitual rate of speech and to pause between each sentence. If they made a reading mistake, they were asked to repeat the whole sentence. Subsequently, participants were asked to read the sentences as fast as they could (at their maximum speed) without becoming unintelligible.

Table 2.4 Sentences for fast and habitual pace

Peter passed the tissues to Ken at the party.
Pauline pushed the cookies over to Tom.
Paul gave tea to Kerry and Tim.
He turned on the TV to watch Kojak.
Can Tina keep the cat for Peter.
Tom calls Paul to help with the parcel.
Tina cooks tuna for Kenny.

Sentences read at different speed can shed light on the ability of the speakers to perform rate changes. Furthermore, it can be tested if a higher degree of vowel reduction is visible in fast read sentences, compared to sentences read at a normal speed. This range in reading pace can also give information on how robust rhythm measures are against articulation rate (Wagner & Dellwo 2004).

2.2.3 Full and Reduced Vowel Sentences

The material described here is taken from Low et al. 2000 (see table 2.5). They were two sets of 5 sentences, one including only full vowels (Full Vowel Set), and the other with an alternation of full and potentially reduced vowels (Reduced Vowel Set). Participants were asked to read the sentences at their habitual speed and if they made a reading mistake to repeat the whole sentence.

Table 2.5 Full and Reduced Vowel Sentences Sets

Full vowel sentences set:

John came back through France last Sunday.

Don seemed quite cross with John last week.

Paul drives past huge towns by highway.

Jane gets four by post each Thursday.

Grace works through huge mounds each Friday.

Reduced vowel sentences set:

John was sick of Fred and Sandy.

Don was across at Jonathan´s.

Paula passed her trial of courage.

Jane has four to last the winter.

Grace was tired of Matthew Freeman

Taken from: Low et al.2000

The sentences were originally developed to capture differences between stress-timed and syllable-timed languages, and thus have the potential to highlight the rhythmic changes described in ataxic dysarthria (see chapter 1). It is hypothesised that English speakers with ataxic dysarthria show less difference between the two sets of sentences than the control speakers, i.e. less vowel reduction, because speakers with ataxic dysarthria are reported to not reduce vowels.

2.2.4 Reading Passage

The passage was adapted from Lowit-Leuschel & Docherty (2001) (table 2.6). Even though the reading passage is prose, it is a dialogue between speakers, including declarative sentences and questions. The passage is phonetically balanced; it therefore includes all consonants of English. All consonants occur in initial position, with the majority also occurring in medial and final position. Furthermore, it includes a selection of consonant clusters. The passage is 179 syllables long.

Participants were asked to read through the passage silently before reading it at their habitual speed out loud. If a reading mistake occurred they were asked to read the whole sentence again.

Table 2.6 Reading Passage

"Good morning Tom."

"Hello, Ken. How are you today? Did you have a good time last night?"

"Yes thanks, we went down town to see a film. Do you have any plans for today yet?"

"Yes, I'll be going to the exhibition in the Todd centre. Would you like to come as well? Pam told my brother yesterday that it is better than she expected. Why don't you meet me there at about two."

"I would with pleasure but I'm not sure whether I can join you. I promised my nephew to take him to the zoo to see the new camels and tigers that came in last week."

"Must you really? You could take him next time round, the exhibition will close soon. And we could visit that posh new pub 'The Cherry Tree' on the way back."

"It will close soon? In that case I'll come. All right."

The current reading passage was preferred for this study instead of the Grandfather Passage which is often used in dysarthric research. The latter is also phonetically balanced, but even though both passages are prose the Grandfather passage (Darley et al. 1975) evokes a slightly different reading style compared to the current passage. The current passage is a dialogue and has intonational variation with direct and indirect speech and some alternation between questions and statements with the aim of eliciting a more naturalistic speech sample than other reading passages.

2.2.5 Spontaneous Speech

Recording entirely spontaneous and natural speech is reported to be an almost impossible task if the speaker is aware of the recording situation (Crystal & Davy, 1969). Speakers are usually inhibited and change the way of speaking when a microphone is visible.

Different approaches can be found to obtain more naturalistic speech, for example by distracting the participant with other materials, e.g. to provide the participant with a *Map Task* (for details see Anderson, Bader, Gurman Bard, Bayle, Doherty, Garrod, Usard, Kowtko, McAllister, Miller, Sotillo & Thomson 1991). Another way of eliciting similar vocabulary from different speakers is to set a certain topic (e.g. a favourite holiday, typical evening at home, family or a restaurant visit). Depending on the chosen topic speakers might become more vocal and caught up in telling, so that the focus is taken from the recording situation. Additionally, the speech output might be closer to the actual habitual spontaneous speech behaviour of the participant and therefore ecologically more valid than a map task.

Therefore, to elicit spontaneous speech which should be relatively comparable across speakers a setting topic was chosen for the current study. The participants were asked to tell the examiner about a restaurant visit, where he/ she went, what they liked to eat, etc. The examiner did not interrupt the participant unless the participant needed prompting to continue talking for longer.

The five different tasks described above comprise speech tasks, which are close to the natural speech behaviour of the participants (i.e. Spontaneous Speech task) and other tasks which are structured and less natural reading tasks. For the latter a predetermined stress pattern is given and they lie on the opposite side of a spectrum of naturalness of speech (i.e. the Poem).

Even though the aim was to collect all data from each participant that took part in the study, it was not possible. Speaker AE_05 did not read the sentences at maximum and habitual speed, because she was too tired to continue recording and no second recording appointment was available. All other speakers took part in all tasks[6].

2.3 Recording location, time and equipment

The participants were recorded at different locations. The participants with ataxic dysarthria were recorded in the clinic they were attending at the time in the USA. The American English control speakers were all recorded at their home or at the Division of Speech and Language Therapy at Strathclyde University in Glasgow. All recordings took place in a quiet room with no other people than the participant and examiner present.

Participants were recorded at a time suitable and convenient to them, avoiding times when they would be tired or distracted. As a result, the times of

[6] See table A in the appendix.

recording varied for the different participants. Since the ataxic dysarthric participants were not on any medication which could influence their speech performance, no particular time of the day was predetermined.

For each speaker the data was recorded in a single recording session, lasting approximately 20 to 30 minutes. Speech was recorded on Digital Audio Tape (DAT) with a DAT recorder (Tascam DA-P1) and a standing microphone (Beyerdynamic Microphone M58) placed on a table in front of the speaker, approximately 40 cm away from the mouth.

To settle the patient into the recording session, the examiner had a short conversation with the participant at the beginning of the recording prior to starting the speech tasks. This was a semi-standardised conversation about the personal life, health, education, family and hobbies of the participant.

2.4 Analysis

The following sections report on the analyses which were carried out on the collected material. The method for the perceptual analysis of the data is followed by an introduction of the programmes used for the analysis and the standards applied for the acoustic analyses. A presentation of the chosen rhythm measures is then followed by information about the statistical analyses used for this experiment.

2.4.1 Perceptual Analysis

Many researchers suggest that a combination of perceptual and instrumental (acoustic) measures is the best analysis technique (Ludlow & Bassich 1983, Weismer 1984, Hartelius et al. 2003). This way research can give a comprehensive picture of the participants' speech. This approach is adopted in the current study.

For the perceptual evaluation of the participants' speech rhythm the current study used category ratings of rhythm on a 5 point scale (for details see below). A category rating was preferred over a magnitude rating, because five point scales are well established in dysarthria assessment (e.g. Lowit 2014) and the listener judges are used to that kind of method, thus increasing chances of good reliability (for details on category ratings see section 1.5 Perceptual Evaluation, p. 39 ff).

The perceptual evaluation was conducted by native speakers of English. Two listener groups of listener judges were recruited. The first group consisted of 4th year honours students (group1 N=9) of the Speech and Language Therapy Division, University of Strathclyde who were not available for further evaluation, since they had already left university. The second group consisted of 2nd year (group2 n=10) students of the same Division. All listener judges were familiar with clinical characteristics of dysarthria, but not with the topic of the study. The students already had some working experience with dysarthric speakers. The first group did not rate the intelligibility of the participants' speech, whereas the second group rated rhythm and intelligibility.

The Poem, Reading Passage, and Spontaneous Speech sample of every participant were perceptually rated for rhythm. All speaker recordings were put in a randomized order within categories, i.e. the listeners heard all recordings of the e.g. Poem for all speakers in a randomized order, followed by the Reading Passage and Spontaneous Speech sample. Two speakers were included twice in each category for rater reliability purposes. One experimental session was held with each group. Before the actual evaluation, a training session was carried out. Three examples of speakers were played to the listeners, one with a severe rhythmic disturbance and dysarthria, one with a mild ataxic dysarthria and mild rhythmic problems and one control speaker. For the evaluation of the participants' speech rhythm, the listener judges were asked to rate the material on a 5 point scale where 1 indicated normal and 5 highly abnormal rhythm. Results from this first evaluation were then gathered and discussed in the group to make sure the procedure was understood.

The perceptual evaluation was carried out to get an evaluation of the perceived rhythmic disturbance of the speech of the ataxic dysarthric speakers in comparison to their healthy control speakers. Furthermore, the results of the perceptual evaluations were used as a reference measure to see whether the rhythm metrics mirror what is perceived as a disturbed speech rhythm.

2.4.2 Acoustic Analysis

In addition to the perceptual analysis an acoustic analysis was carried out. The equipment, settings and analysis tools are now presented, followed by an argumentation for the chosen rhythm metrics and their introduction. Since the rhythm metrics are based on different measurement units, an explanation is given about how the different units are labelled and also on which task material each metric was applied.

2.4.2.1 Equipment, Settings & Praat (Text Grids & Scripts)

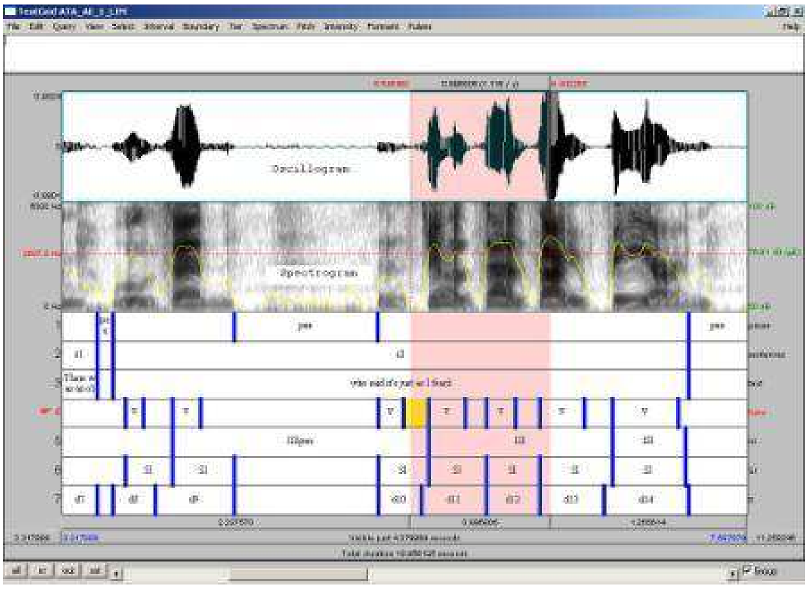
Data was captured from the DAT recorder with Multi Speech at a sampling rate of 44 kHz. Acoustic analysis was performed with Praat (Version 4.4.24) (Boersma & Weenink 2013) and Multi Speech (Version 2.7).

The oscillograms and wide-band spectrograms of the data were displayed on the monitor screen; and pauses, vowels, syllables and inter-stress-intervals were labelled in Praat TextGrids.

The calculation of the various rhythm measures was automated with Praat scripts. The measurement units, such as vowels, syllables, etc. were hand labelled and saved as a TextGrid file together with the acoustic signal (see figure 2.1). Scripts for each rhythm measure were then run with this file. Information about the boundaries of the measured units can be found in the following section. Labels were put at the beginning and end of a unit/ interval on one tier, the beginning and end of pauses on a different tier.

Scripts were run for extracting unit durations, as well as for calculating rhythm measures. To check if the scripts were working and calculating correctly, files were analysed manually and results compared, i.e. the percentage of agreement between the manual calculations and calculations conducted with Praat scripts was determined. The agreement was very high (100%).[7]

Figure 2.1 Screenshot of Praat Edit Window



2.4.3 Rhythm Metrics

In the following section a detailed description of the rhythm metrics applied to the collected speech data is given. For the current study the normalised Pairwise Variability Index (nPVI, Low et al. 2000), the measure of vocalic intervals (%V, Ramus et al. 1999), the VarcoV (White & Mattys 2007 a,b), the Variability Index (VI, Deterding 2001), and the Inter-Stress Interval measure (ISI,

[7] The nPVI was furthermore controlled via an Excel spreadsheet which is available for download from Esther Grabe's homepage and includes the nPVI calculations. (http://www.phon.ox.ac.uk/files/people/grabe/ 30.03.2010)

Hartelius et al. 2000) have been chosen. An overview of the measures and the materials they were applied is given in table 2.7.

Table 2.7 Rhythm metrics applied to material

Rhythm Metric	Materials
nPVI	Poem, Reading Passage, Spontaneous Speech, Sentences fast & habitual pace, Reduced & Full Vowel Sentences
%V	
VarcoV	
VI	Poem
ISI	

The three vowel based metrics nPVI, VarcoV and %V have been applied to all speech tasks. Additionally, the VI and ISI metric have also been applied to the Poem. All metrics have not been applied to all tasks, as the in the Poem interstress-intervals and syllables could be identified most reliably. The vowel based metrics, however, have been applied to all tasks, as vowels can be reliably identified across different speech tasks.

2.4.3.1 Normalised Pairwise Variability Index (nPVI)

The nPVI (Low et al. 2000) is compiled by calculating the difference in duration of successive vowels, and dividing the absolute value of this difference by the mean duration of the successive vowel pairs (see table 2.7). This is then summed up and divided by the number of pairs. A multiplication by 100 is carried out because of fractional values acquired by the normalisation procedure for articulation rate. The normalisation carried out with the nPVI is a local normalization since a division of the mean duration of each vowel pair is calculated. As opposed to the VI, in the nPVI the normalisation is part of the formula (division by the mean vowel duration) (see figure 2.2, and table 2.7 for the complete nPVI formula).

Figure 2.2 Normalisation included in the nPVI

$$\frac{1}{\left(x_{(i)} + x_{(i+1)}\right)/2}$$

High nPVI values mirror a high durational variability between successive vowels, i.e. a tendency towards stress-timing; whereas low nPVI values reflect little durational variability between successive vowels (a tendency towards syllable-timing). For the current study an average of the nPVIs across all sentences within each task material was taken and used for a comparison of the speakers. Pre-pausal and utterance final intervals were not excluded following Grabe & Low (2002), even though a possibility of lengthening effects exists. Phrase-final lengthening may appear in the absence of pausing and not all pausing may be preceded by lengthening, e.g. hesitation or disfluent pauses.

Furthermore, "the locus and extent of final lengthening and other prosodic lengthening processes (…) may be language specific and may contribute to the overall perception of cross-linguistic differences in rhythmicity" (White & Mattys 2007a, p.507). Thus the inclusion of the pre-pausal and utterance final intervals may lead to a comprehensive picture of the (perceived) rhythmic nature of the speech tasks. This may also count for the perception of the differences in rhythmicity between the speech of ataxic dysarthric and control speakers, as e.g. prosodic lengthening effects may occur in ataxic dysarthric speech.

The nPVI was chosen for the current study, as it is not only able to distinguish between stress- and syllable-timed languages, but it has also been applied to ataxic dysarthric speech where it was found to reflect rhythmic differences between ataxic and control speakers (Stuntebeck 2002). Additionally, the nPVI was chosen, instead of later proposed metrics, as it has not yet been fully investigated how it reflects the perceived impression of the speech disturbance of ataxic dysarthric speakers. Later studies proposed variants of the nPVI such as the rPVI-C (White & Mattys 2007a) or nPVI-VC (Liss et al. 2009), differently measuring the consonantal or intervocalic and not the vocalic intervals and not including a normalisation. However, preference is given to the nPVI, as it is not only able to distinguish between stress- and syllable- timed languages, but also showed differences between ataxic dysarthric and control speakers (Stuntebeck 2002, Liss et al. 2009).

2.4.3.2 Percentage of vowel duration (%V)

The %V metric is a measure of the total duration of vocalic intervals within an utterance in relation to consonantal segments and does not consider the sequence of intervals (Ramus et al. 1999). The sum of the duration of all vocalic intervals is taken, with a vocalic interval being defined as the interval between the beginning and end of a vowel, or a cluster of vowels. The percentage of vocalic intervals (%V) is the sum of vocalic intervals divided by the total duration of the sentences. A high %V value indicates a tendency towards syllable-timed rhythm.

Ramus et al. (1999) proposed several rhythm metrics (see 1.4.2.1) of which the %V was chosen for the current study, due to several reasons: First the authors found the %V measure as being the best acoustic correlate of rhythm class. It furthermore can be easily applied to a range of different tasks, as vowels can generally be reliably identified by a trained phonetician. Even though the %V metric does not include a normalisation for articulation rate, the comparison of different metrics with and without an included normalisation can shed light on the importance of it.

2.4.3.3 VarcoV

The rate normalised rhythm measure VarcoV (White & Mattys 2007 a,b) is an interval measure based on vowel durations. To calculate the VarcoV the standard deviation of the vocalic interval is divided by the mean vocalic interval duration and multiplied by 100. For vowel measurements, sentence final vowels, as well as for pausing the same guidelines were applied as for the nPVI and %V.

2.4.3.4 Variability Index

The Variability Index (VI) (Deterding 1994, Deterding 2001) is a syllable based metric, where the difference in duration of neighbouring syllables are calculated and a normalisation across the utterance is applied with the exclusion of the last syllable (see also chapter 1.4.2.3 for details). The author excluded the final syllable from their analysis, because of possible final-syllable lengthening effects.

To take possible effects of speaking rate into consideration a normalisation was included in the VI, where the duration of each syllable was divided by the average duration of all syllables without the final syllable (Deterding 2001:222).

For the current study a range of different rhythm metrics were chosen which were based on different measurement units. Differently to the above mentioned rhythm metrics the VI is based on syllable durations and not vowel based. As the study aims to test a range of different rhythm metrics the VI was included.

2.4.3.5 Coefficient of Variation of Interstress Interval Durations (ISI)

For the ISI the coefficient of variation (CV%) of the interstress interval durations was calculated. The ISI is based on the duration of stress intervals or feet. The Poem consisted of 13 main stresses; therefore 12 Inter-Stress-Intervals were measured (see table 2.8). According to Fant, Kruckenberg & Nord (1989) boundary spanning feet were excluded from the calculations. A boundary spanning foot can be found when the speakers pauses within a foot. Therefore the number of feet differed across speakers.

Table 2.8 Inter-Stress Intervals in the Poem

There w/as an old m/an with a b/eard

1 2 3

Who s/aid it's j/ust as I f/eared

4 5 6

Two /owls and a h/en

7 8

Four l/arks and a wr/en

9 10

Have /all build their n/est in my b/eard

11 12

2.4.4 Measurement units

The location of the vowel-consonant or consonant-vowel boundaries, as well as syllable boundaries or Inter-stress-interval boundaries were identified and labelled. This was primarily done by visual examination of the wideband spectrograms and speech waveforms in Praat (Boersma & Weenink 2013).

Vowels were measured where clearly visible vocalic formant structure in the spectrogram could be seen. The vowel onset was marked at the onset of this formant structure and the first regular vertical striations of the vowel and the onset of the second formant respectively. The offset was indicated at the ending of the energy in F2 or higher formants. Furthermore, the acoustic analysis was based on criteria from Peterson & Lehiste (1960).

Problems could occur with the presence or absence of vocalic segments, and the devoicing of vocalic segments. The specifications of Low et al. (2000) were followed: If after a voiceless plosive the vowel could not be seen in the spectrogram/ oscillogram the duration of the plosive preceding the vowel was measured as the vowel duration. The same rule was applied in case of the syllabification of a consonant. The durations of the vocalic intervals were taken, excluding pauses but across sentence boundaries (according to Low et al. 2000).

Syllable boundaries were annotated for the VI. For the determination of syllables analysed the procedure of Deterding (2001) is followed: In cases of merging words ("we are" - "we're"), triphthongs (e.g. "hire" or "hour" as two syllables or one) and coarticulatiory effects due to speaking rate (e.g. "comfortable" as a three or four syllable word), always the smaller number of syllables was assumed (Deterding 2001:221). In addition, the syllable durations were measured following the *maximal onset principle (MOP)* which assigns intervocalic consonants to the succeeding syllable as long as it is in conformity with the phonotactic constraints of syllable onsets, e.g. the word *diploma* can be divided in several ways: *dip.lo.ma* vs. *di.plo.ma* (Wagner & Dellwo 2004). However, the only division that is in conformity with the maximal onset principle is *di.plo.ma.*

Inter-Stress-Intervals were defined as starting with a stressed vowel and ending with the consonant before the next stressed vowel (Hartelius et al. 2000). Intervals including a pause (boundary spanning foot) were not included in the calculation in line with Fant et al. (1989).

Table 2.9 Inter-stress Intervals in the Poem

	Inter-stress Interval	
	There	
w	as an old m	1
	an with a b	2
	eard.	3
Who s	aid it's j	4
	ust as I f	5
	eard.	6
Two	owls and a h	7
	en.	8
four l	arks and a wr	9
	en	10
Have	all build their n	11
	est in my b	12
	eard	(13)

Silent pauses were marked if the intensity of the signal was 1.5 standard deviations below the mean intensity of all voiced intervals (MIV) or smaller. Minimum pause length was set at 250ms, except when there were clear intonational breaks or breath pauses. These were also labelled as pauses and then excluded from the measured interval (Low et al. 2000).

2.4.5 Articulation rate

Research shows that some rhythm measures were related to speech rate (Barry & Russo 2003, Dellwo & Wagner 2003, Wagner & Dellwo 2004, White & Mattys 2007a, Liss et al. 2009). In studies with ataxic dysarthric speech compared to control speech, it is especially important that the rhythm metrics are not related to rate, since the speakers with ataxic dysarthria often present a much lower speech and articulation rate compared to the control speakers (Kent et al. 1997, Cannito & Marquardt 1997, Duffy 2012, Brown et al.1970). Articulation rate was taken for the current study as opposed to speech rate, as the rhythm measures were also calculated without pauses.

The articulation rate per speaker per task was calculated (syllables per minute without pausing time).

2.5 Statistical Analysis

For the statistical analysis SPSS 16.0 statistical package was used. All participants were divided into two groups, one with speakers with ataxic dysarthria, and the other with healthy control speakers. Group size was relatively small with six speakers per group. However, appropriate measures were selected to consider this.

Non-parametric tests were chosen for this study. Non-parametric tests do not ask for a normal distribution of scores, as do parametric tests. A normal distribution may not be given in speakers with ataxic dysarthria due to variability of performance in these speakers.

As a measure of the strength of the relationship between the results of the perceptual evaluation of the participants' speech and each rhythm metric, as well as an evaluation of the validity of the acoustic measures Spearman correlations were carried out. Spearman correlations give information about the direction and strength of the relationship between two variables. It was also calculated between the results for the articulation rate and rhythm metrics.

The conventional test for exploring differences between groups is the T-test. However, the non-parametric alternative the Mann-Whitney-U-Test (for two independent samples) was used to test for group differences. The Mann-Whitney-U-Test does not compare single values, but allocates ranks to each score independent of the group.

To compare the performance of one group in different tasks the Wilcoxon Signed Rank Test was carried out. It considers information about the significance of magnitude of difference between pairs, as well as the difference between the pairs. Significance levels were at $p = 0.05$.

Bonferroni corrections were not carried out as of the exploratory nature of the experiment.

Table 2.10 Overview of the statistical analysis applied to the data

Comparison	Test
Group differences	Mann-Whitney-U-Test
Group differences in tasks	Wilcoxon Signed Rank Test
Task differences	Wilcoxon Signed Rank Test
Correlation between the results of the perceptual evaluation of the participants' speech & each rhythm metric	Spearman Correlation
Correlation between the results of the articulation rate and rhythm metrics	Spearman Correlation

2.5.1 Inter- and Intra-rater reliability

To determine intra- and inter-rater reliability of the set labels and rhythm metrics data of randomly chosen speakers were relabelled and remeasured by the same and another examiner using Praat (approximately 8% of the data). The vocalic intervals, as well as syllables, and inter-stress intervals were re-labelled and durations compared. The Cronbach's alpha – a coefficient of reliability (or internal consistency) – was calculated for inter- and intra-rater reliability with regard to the acoustic analysis. Cronbach's alpha score for intra-rater reliability is .986 and ranged between .956 - .968 for inter-rater reliability. The scores were taken as good (George & Mallery, 2002).

The inter- and intra-rater reliability of the listener judges of the perceptual evaluation was tested with the Kendalls W. The Kendalls W is similar to the Cronbach's alpha, but it is the appropriate measure for multiple spot tests. Is also takes into consideration the actual number of the evaluation the listener judges chose (one to five on a scale, where one indicated normal and 5 highly abnormal rhythm).

The listener judges scores for the group of speakers with ataxic dysarthria are more variably than the for the control group. However, listener judges showed similar tendencies in variability. Statistic analysis with the Kendalls W shows good agreement between the listener judges for all three speech task that were rated (Poem W= 0.784 p= 0.000, Reading Passage W= 0.823 p= 0.000, Spontaneous Speech W= 0.832 p= 0.000).

Additionally, speech samples of two speakers of each speech task were randomly included twice in the perceptual evaluation and therefore rated twice.

2.6 Summary

This chapter presented the methodologies applied in this study. In the first part of this chapter, the six ataxic dysarthric and control participants were introduced. This was followed by a detailed description of the seven tasks (Poem, Reading Passage, Spontaneous Speech sample, Fast Rate Sentences, Habitual Rate Sentences, Full Vowel Sentences, Reduced Vowel Sentences) and the five rhythm metrics (nPVI, %V, VarcoV, VI, ISI). In a next step, the data collection and analysis procedures were outlined, including the acoustic analysis and the perceptual evaluation of the data. Finally, the statistical procedures were introduced and the results of the reliability measures presented.

As indicated above, the study was conducted to test which rhythm metric and material best differentiate between speakers with ataxic dysarthria and controls and correlated with the perceptual evaluation of the participants' speech rhythm. The findings are presented in the next chapter, followed by a discussion and conclusion.

3. Results

Based on empirical findings and theoretical assumptions laid out in the previous chapters, the aims of this study were to see which rhythm metric and task would be most suited to differentiate ataxic dysarthric speech as compared to the speech of healthy control speakers, and what is the correlation between rhythm metric results and the and the perceptual evaluation of the participants' speech rhythm. Rhythm metric results were furthermore correlated with articulation rate.

As described in chapter 2, the participants took part in seven different tasks, i.e. Poem, Reading Passage, Spontaneous Speech, Full Vowel Sentences, Reduced Vowel Sentences, as well as Habitual and Fast Rate Sentences. Altogether, five different rhythm metrics were used in the study and applied to one task, the Poem. Three of the metrics were additionally applied to the other six tasks (nPVI, VarcoV, and %V).

The results are presented in four sections: In the first section, the results for the rhythm metrics which have been applied across all tasks are presented (nPVI, VarcoV and %V). The results for the additional rhythm measures which have been applied to the Poem task only (ISI, VI) are subsequently presented in the second section.

In section three, results of the perceptual evaluation and their correlation with the acoustic rhythm measures are presented.

The fourth section of this chapter reports the results of the articulation rate measures and their correlation with the results of the rhythm measures.

Group and task differences have been analysed for all measures. Graphs and tables are used to highlight results; group mean values, standard deviations. Tables and graphs with all individual values for all speakers, tasks and rhythm metrics can be found in the appendix.

3.1 Results for the nPVI, VarcoV and %V for all tasks

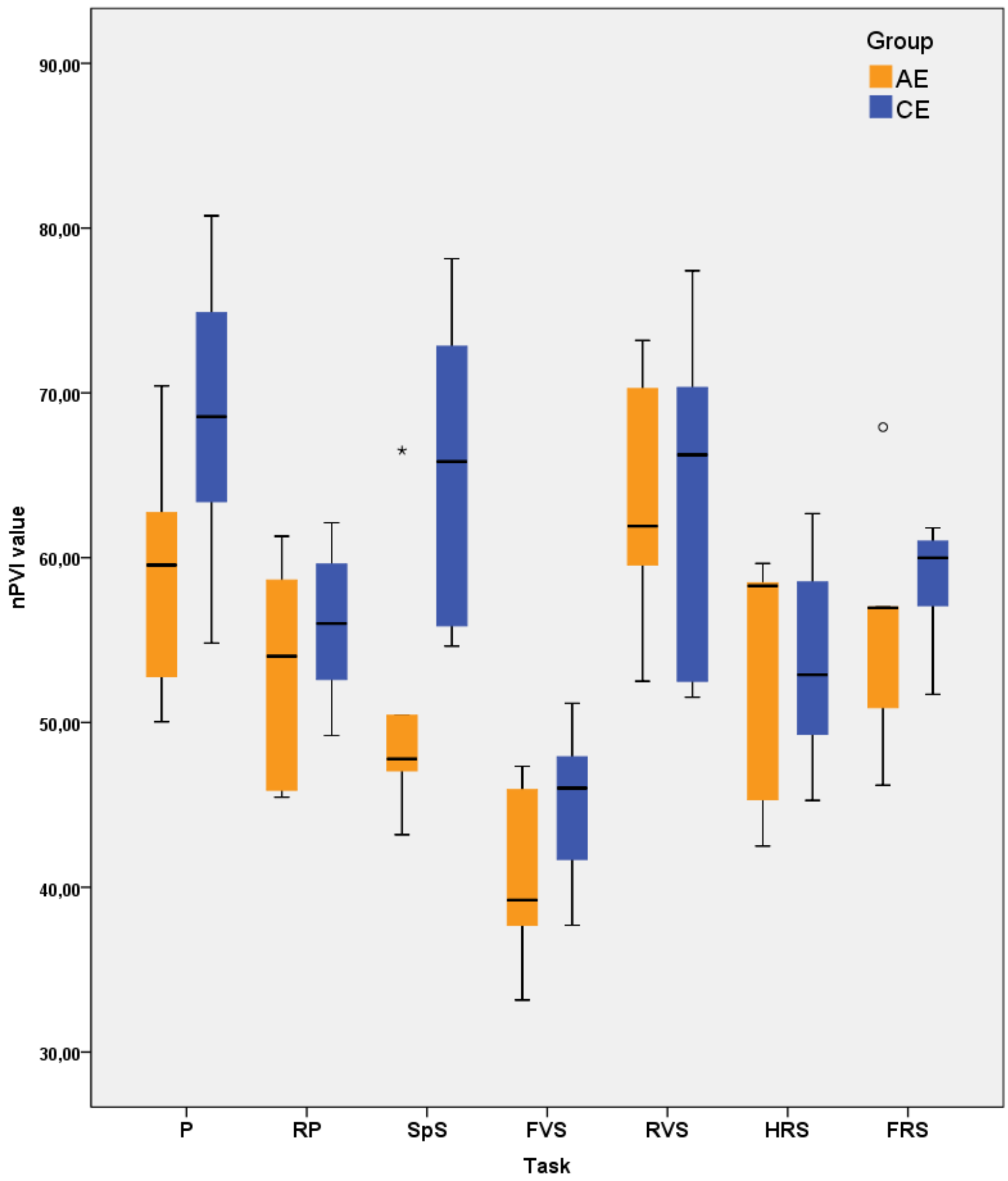
3.1.1 nPVI

The normalised Pairwise Variability Index (nPVI) is a rhythm measure based on the calculation of successive vowel durations. A low nPVI value indicates little variability in vowel length, which is supposed to reflect a more syllable-timed rhythm; a high nPVI value indicates higher variability in successive vowel length, i.e. a more stress-timed.

Group differences

The following figure shows the mean (the 50^{th} percentile as the line in the middle), range (the whiskers on the top and bottom of the box), the 75^{th} and 25^{th} percentile (the box above and below the line), as well as outliers (circles or stars above the whiskers) for the ataxic dysarthric and control group for the nPVI for all tasks.

Figure 3.1 Mean nPVI values and range for all tasks



Abbreviations: AE= English speaking ataxic dysarthric group, CE= English speaking control group, P = Poem, RP= Reading Passage, SpS= Spontaneous Speech, FVS = Full Vowel Sentences, RVS = Reduced Vowel Sentences, HRS = Habitual Rate Sentences, FRS = Fast Rate Sentences

Table 3.1 Differences between the ataxic dysarthric and control group for each task - nPVI (Mann-Whitney-U-Test)

P	RP	SpS	FVS	RVS	HRS	FRS
p= 0.078 Z= -1.76	P= 0.522 Z= -0.641	p= 0.016* Z= -2.40	p= 0.128 Z= -1.52	p= 0.810 Z= -0.24	P= 0.855 Z= -0.18	p= 0.201 Z= -1.27
* difference is significant at the 0.05 level, Abbreviations: P = Poem, RP= Reading Passage, SpS= Spontaneous Speech, FVS = Full Vowel Sentences, RVS = Reduced Vowel Sentences, HRS = Habitual Rate Sentences, FRS = Fast Rate Sentences						

Figure 3.1 shows that group ranges overlap considerably (see also group means and standard deviations in appendix B-E). Consequently, most group comparisons yielded non-significant results, with only the Spontaneous Speech task differentiating the two groups from each other (table 3.1). Having said that, the data for the Poem the Fast Rate Sentences also show a tendency of higher values, i.e. more stress-timed rhythm, for the control speakers compared to the ataxic dysarthric group. The remaining tasks show very little difference between groups.

Task differences

Statistical analysis with the Wilcoxon-Signed-Rank-Test showed some significant differences between the tasks. Table 3.2 shows the comparisons between the Poem, Reading Passage and Spontaneous Speech; the Full Vowel Sentences with the Reduced Vowel Sentences; and the Normal Rate Sentences with the Fast Rate Sentences.

Table 3.2 Significant differences between the tasks for both groups - nPVI (Wilcoxon Signed Rank Test)

		Ataxic dysarthric	Control
Poem	Reading Passage	p= 0.028* Z= -2.20	-
Poem	Spontaneous Speech	p= 0.028* Z= -2.20	-
Reading Passage	Spontaneous Speech	-	p= 0.046* Z=-1.99
Full Vowel Sentences	Reduced Vowel Sentences	p= 0.028* Z= -2.20	p= 0.028* Z= -2.20
Habitual Rate Sentences	Fast Rate Sentences	-	-
*difference is significant at the 0.05 level			

For the ataxic dysarthric group the Poem differed significantly from the Reading Passage and the Spontaneous Speech sample, but not for the control group. For the latter group, however, the Reading Passage differed significantly from the Spontaneous Speech sample. The Full Vowel Sentences differ from Reduced Vowel Sentences for both groups.

While the nPVI values differ across the tasks, Full Vowel Sentences and Reduced Vowel Sentences/ Poem seem to be at either side of the continuum for the ataxic dysarthric group. Full Vowel Sentences also present the lowest values for the control group, but the highest can be found for the Poem and Spontaneous Speech task. The high nPVI values in the Poem, Reduced Vowel

Sentences, and Spontaneous Speech sample represent a more stress-timed rhythm compared to the low valued Full Vowel Sentences.

Summarizing the results for the nPVI, one significant group difference was found for the Spontaneous Speech task. Significant differences between the tasks were apparent for both groups in several different tasks.

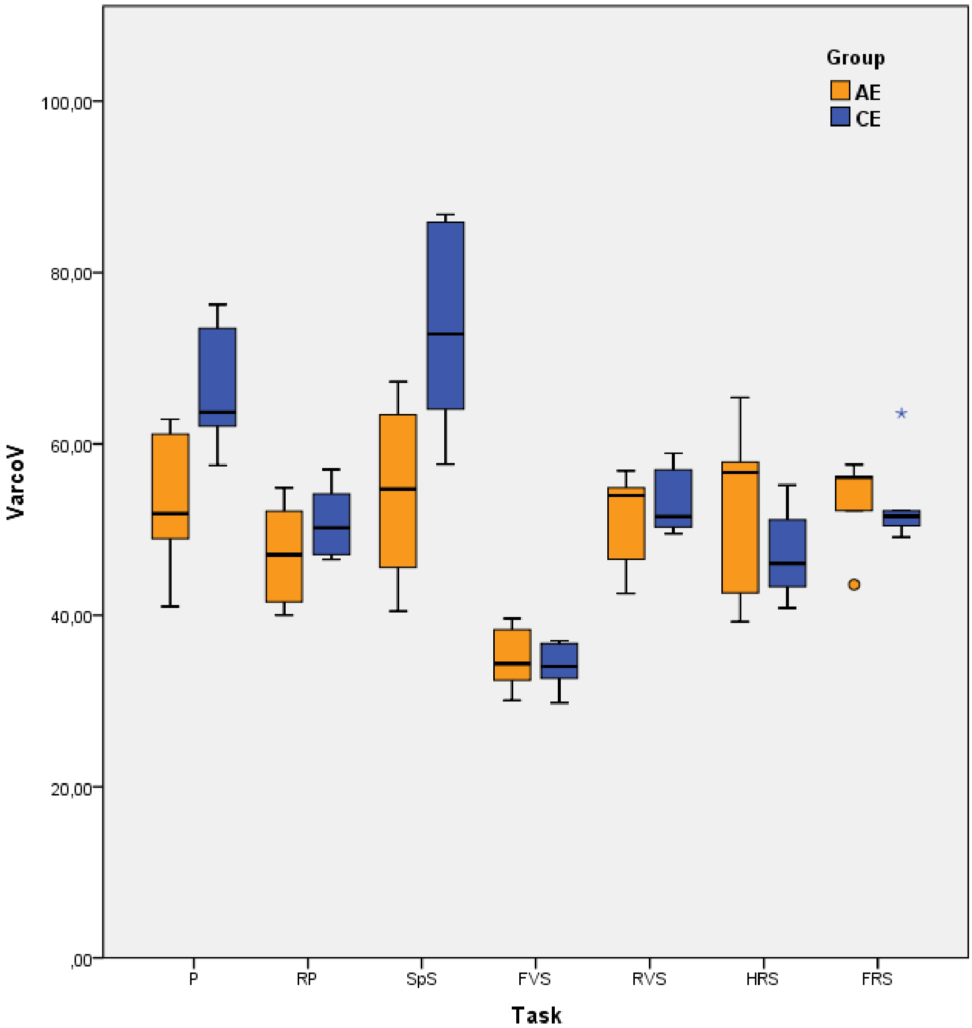
3.1.2 VarcoV

The VarcoV measure is a rate-normalised interval measure based on vowel durations. A high VarcoV value indicates a stress-timed rhythm and a low syllable-timed rhythm.

Group Differences

The means and the range for all tasks are shown in figure 3.2 and suggest difference in performance between the groups for the Poem and Spontaneous Speech, with less variation for the remaining tasks. Where differences are shown, the control group performs in the more stressed timed range again, with the exception of the Fast Rate Sentences where this relationship is reversed. The qualitative evaluation is largely confirmed by the Mann-Whitney U Test which yielded significant results for the Poem and Spontaneous Speech although not for the Fast Rate Sentences (see table 3.3).

Figure 3.2 Mean VarcoV values and range for all tasks



Abbreviations: AE= English speaking ataxic dysarthric group, CE= English speaking control group, P = Poem, RP = Reading Passage, SpS = Spontaneous Speech, FVS = Full Vowel Sentences, RVS = Reduced Vowel Sentences, HRS = Habitual Rate Sentences, FRS = Fast Rate Sentences

Table 3.3 Differences between the ataxic dysarthric and control group for each task – VarcoV (Mann-Whitney-Test)

P	RP	SpS	FVS	RVS	HRS	FRS
p= 0.016* Z= -2.40	p= 0.423 Z= -0.80	P= 0.025* Z= -2.24	p= 0.749 Z= -0.32	p= 0.749 Z= -0.32	p= 0.465 Z= -0.73	p= 0.361 Z= -0.91

* difference is significant at the 0.05 level, Abbreviations: P = Poem, RP= Reading Passage, SpS= Spontaneous Speech, FVS = Full Vowel Sentences, RVS = Reduced Vowel Sentences, HRS = Habitual Rate Sentences, FRS = Fast Rate Sentences

Task differences

The Wilcoxon-Signed-Rank-Test yielded three significant differences between the tasks for the ataxic dysarthric and four for the control group (Table 3.4). For both groups VarcoV values for the Reading Passage are lower (indicating a more syllable-timed rhythm) compared to the Poem and Spontaneous Speech Sample. Again for both groups, the Full vowel sentences show the lowest VarcoV values of alls tasks, whereas the Spontaneous Speech sample shows the most stress-timed rhythm. The control group showed an added difference between Habitual and Fast Rate Sentences.

Table 3.4 Significant differences between the tasks for both groups - VarcoV (Wilcoxon Signed Rank Test)

		Ataxic dysarthric	Control
Poem	Reading Passage	p= 0.046* Z= -1.99	p= 0. 028* Z= -2.20
Poem	Spontaneous Speech	-	-
Reading Passage	Spontaneous Speech	p= 0.046* Z= -1.99	p= 0. 028* Z= -2.20
Full Vowel Sent.	Reduced Vowel Sent.	p= 0.028* Z= -2.20	p= 0. 028* Z= -2.20
Habitual Pace Sent.	Fast Pace Sent	-	p= 0. 028* Z= -2.20
*difference is significant at the 0.05 level			

In summary, the VarcoV measure showed significant group differences for the Poem and Spontaneous Speech tasks. Furthermore several differences between the tasks were found to be significant for both groups. Compared to the nPVI, the VarcoV showed more groups differences, as well as more task differences for both groups.

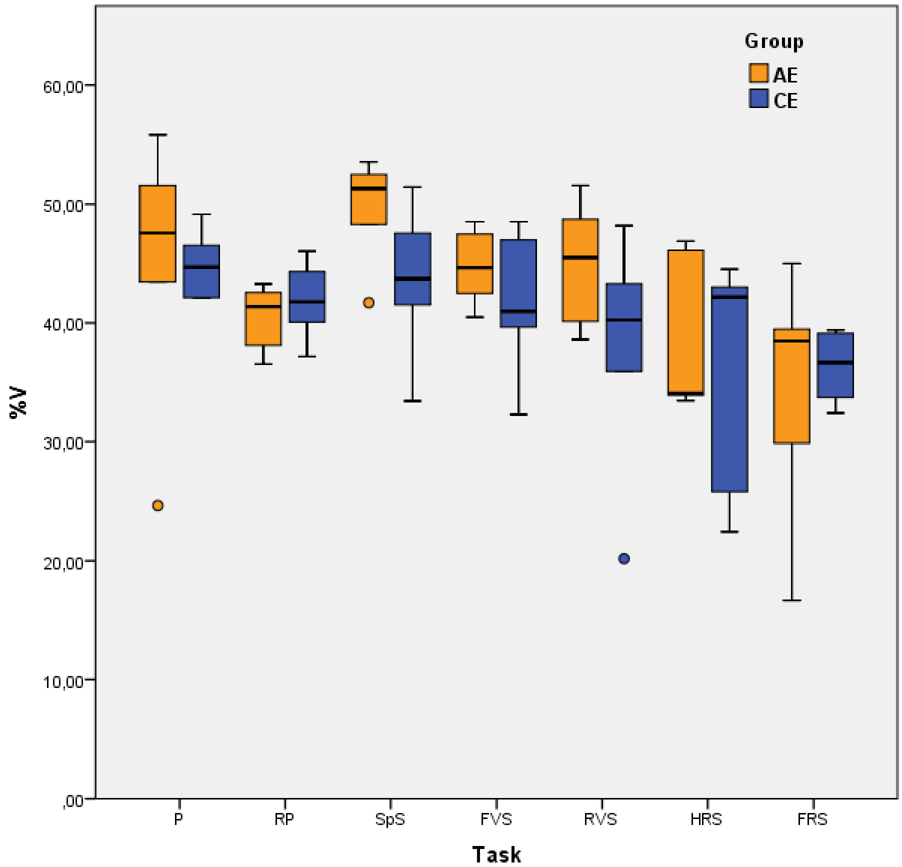
3.1.3 %V

As proposed by Ramus et al. (1999) the %V metric measures the percentage of vowel interval duration within an utterance. In contrast to the nPVI and Varco V a high %V value indicates a syllable-timed rhythm and a low %V value a stress-timed rhythm.

Group differences

Figure 3.3 shows the means and range for all tasks for the %V measure. The Mann-Whitney U Test results reveal no significant difference between the two groups for any of the tasks (table 3.5). Only the Spontaneous Speech data suggest a difference between the groups, however, this is rendered non-significant, although close to the significance threshold of 5% by the performance of an outlier (P= 0.065, Z= -1.92). Despite the lack of significant results, the group mean and range show a tendency for the control group to show lower %V values, and thus a more stress-timed rhythm, with the exception of the Fast Rate Sentences and the Reading Passage.

Figure 3.3 Mean %V values and range for all tasks



Abbreviations: AE= English speaking ataxic dysarthric group, CE= English speaking control group, P = Poem, RP = Reading Passage, SpS = Spontaneous Speech, FVS = Full Vowel Sentences, RVS = Reduced Vowel Sentences, HRS = Habitual Rate Sentences, FRS = Fast Rate Sentences

Table 3.5 Differences between the ataxic dysarthric and control group for each task - %V (Mann-Whitney-Test)

P	RP	SpS	FVS	RVS	HRS	FRS
p= 0.337 Z= -0.096	p= 0.522 Z= -0.64	P= 0.065 Z= -1.92	p= 0.297 Z= -1.04	p= 0.631 Z= -0.48	p= 0.584 Z= -0.55	P= 0.855 Z= -0.18
* difference is significant at the 0.05 level, Abbreviations: Reduced V. Sentences=Reduced Vowel Sentences						

Task differences

The %V yielded very few significant results for the task comparison with only one significant difference between the Reading Passage and Spontaneous Speech for the ataxic dysarthric group, and none at all for the control group (table 3.6).

Table 3.6 Significant differences between the tasks for both groups - %V (Wilcoxon Signed Rank Test)

		Ataxic dysarthric	Control
Poem	Reading Passage		
Poem	Spontaneous Speech		
Reading Passage	Spontaneous Speech	$p = 0.028$* $Z = -2.20$	
Full Vowel Sent.	Reduced Vowel Sent		
Normal Rate Sent	Fast Rate Sent		
*difference is significant at the 0.05 level, Abbreviations: Full Vowel Sent. = Full Vowel Sentences, Reduced Vowel Sent. = Reduced Vowel Sentences, Normal R. Sent. = Normal Rate Sentences, Fast Rate Sent. = Fast Rate Sentences			

Since all %V values lie within a small range, the picture of which task shows the most stress-timed or syllable-timed values is less clear than for the other measures and a trend can not be identified.

In summary, the %V measure did neither yield a significant group difference nor were there many differences between the tasks.

Summary for the nPVI, VarcoV and %V

The nPVI showed only one significant group difference for Spontaneous Speech, as did the VarcoV. Additionally the VarcoV showed a significant group difference for the Poem. The %V failed to produce any significant group differences (even though the result was almost significant for Spontaneous Speech). The ataxic dysarthric group did not always show the highest %V values (indicating a syllable-timed rhythm). The nPVI and VarcoV, however, showed more stress-timed values for the control group.

A difference in performance across the tasks was displayed by some significant task differences for both groups in all three measures. The nPVI and VarcoV, however, highlighted more significant differences between tasks than the %V.

3.2 Results for the ISI and VI measure and the comparison of all measures for the Poem

The following section will first present the results of the four further rhythm measures ISI and VI for the Poem, followed by the comparison of all rhythm measures for this task.

3.2.1 ISI

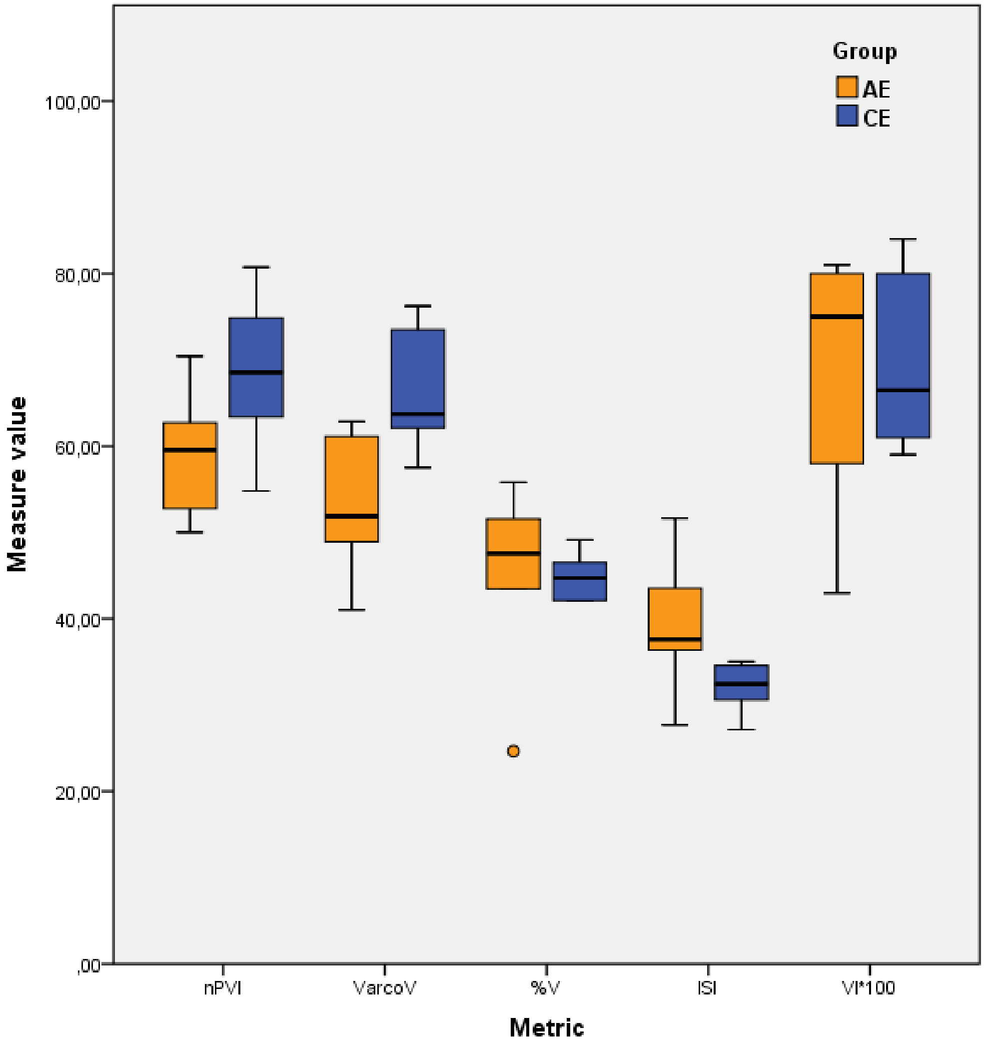
The ISI is the measure of the coefficient of variation across inter-stress intervals. A high ISI measure indicates a syllable-timed rhythm, whereas a low value indicates a stress-timed rhythm. Figure 3.4 shows that the mean value for the control group is lower compared to the ataxic dysarthric group, therefore indicating a more stress-timed rhythm for the control speakers. The ataxic dysarthric group displays a bigger range of values than the control group. The Mann-Whitney U Test revealed a significant difference between the ataxic dysarthric and control group for the ISI (table 3.7).

Table 3.7 Differences between the ataxic dysarthric and control group in the Poem for the ISI & VI (Mann-Whitney-Test)

ISI	VI
p= 0.037* Z= -2.08	P= 0.936 Z= -0.16
* significant at the 0.05 level	

ISI = Inter-Stress Intervals, VI = Variability Index

Figure 3.4 Means and range for the Poem for all measures



Abbreviations: AE= English speaking ataxic dysarthric group, CE= English speaking control group, ISI = Inter-Stress Interval, VI = Variability Index (the VI values were multiplied by 100, for a better comparison), nPVI = normalized Pairwise Variability Index, High value indicates stress-timed rhythm for nPVI, VarcoV & VI, low value indicates stress-timed rhythm for %V & ISI and vice versa.

3.2.2 VI

Syllable-timing is represented by a small value and stress-timing by a high value with the Variability Index. The means and range of both groups for the VI are very alike (figure 3.4). A high degree of overlap is visible and no significant difference between the groups was evident (table 3.7). [8]

3.2.3 Comparison of all rhythm measures for the Poem

Figure 3.4 shows all means and the range for all rhythm measures for the Poem. Only two rhythm measures (VarcoV and ISI) revealed significant differences between the groups, but a general trend of a more stress-timed rhythm can be seen for all measures but VI.

Summary for the results for ISI & VI and the comparison of all rhythm measures for the Poem

Out of the five rhythm metrics, the ISI and VarcoV showed significant differences between the groups for the Poem. Whereas the range was small for the ISI and no overlap between the groups was obvious, a high degree of overlap and a high range was displayed by the VI. A general trend of more stress-timed values for the control group could be seen for most of the measures except the VI.

---

[8] The values of the Variability Index (VI) were multiplied by 100 for display and comparison purposes.

3.3 Perceptual Evaluation of the participants' rhythm and correlation between the perceptual evaluation and rhythm metrics

This section presents the results of the perceptual evaluation of the participants' speech, followed by the correlation between the perceptual evaluation and the results of the rhythm measures.

3.3.1 Perceptual Evaluation of the participants' rhythm

The participants' speech was rated with regard to rhythm by a group of Speech and Language Therapy students for three different tasks (Poem, Reading Passage, and Spontaneous Speech), to get an evaluation of the perceived rhythmic disturbance of the ataxic dysarthric speakers in comparison to their healthy control speaker.

The speech rhythm, especially the presence or absence of a rhythmic disturbance, was rated for all participants on a scale from one to five, where one represents a normal and five an abnormal rhythm.

The Mann-Whithney-U Test showed significant differences between the groups for all three investigated tasks: Poem p= 0.017*, Z= -2.39; Reading Passage p=0.028*, Z= -2.20; Spontaneous Speech p=0.010*, Z=-2.58 (difference significant at the 0.05 level), indicating that the rhythm of the ataxic dysarthric group was perceived as being more disturbed than that of the control group (see also table 3.8).

Table 3.8 Perceptual Evaluation of rhythm

	Poem		Reading Passage		Spontaneous Speech	
Speaker	AE	CE	AE	CE	AE	CE
01	3.06	2.00	3.78	2.67	3.78	2.22
02	2.24	1.56	2.33	1.89	1.11	1.11
03	1.39	1.11	1.67	2.00	1.89	1.33
04	2.83	1.11	3.11	1.00	2.89	1.11
05	3.61	1.00	3.00	1.11	2.89	1.11
06	2.33	1.11	3.00	1.11	2.67	1.56
Mean	2.58	1.31	2.81	1.63	2.54	1.41
SD	0.77	0.35	0.61	0.67	0.92	0.40

Abbreviation: AE= English speaking ataxic dysarthric group, CE= English speaking control group, SD= Standard Deviation, 1 indicates normal, 5 highly abnormal rhythm

3.3.2 Correlation between the perceptual evaluation of rhythm and the rhythm measures

Results of the perceptual evaluation of the participants' rhythm were correlated with the results of the rhythm measures, to see how the perceptual and acoustic measures compared.

Bivariate correlations showed significant correlations between the two analyses (table 3.9) for the nPVI for the Poem and Spontaneous Speech, VarcoV for the Poem, Spontaneous Speech and Reading Passage and the ISI for the Poem. None of the other rhythm measures showed significant results.

Table 3.9 Correlation of the perceptual evaluation of rhythm and rhythm measures

Correlated metrics	Task		
	Poem	Reading Passage	Spontaneous Speech
nPVI & PE	r= -.754** p= 0.005	r= .105 p = 0.105	r= -.685* p= 0.014
VarcoV & PE	r= -.796** p= 0.002	r= -.628* p= 0.029	r= -.624* p= 0.030
%V & PE	r= .373 p =0.232	r= -.221 p= 0.490	r= .510 p= 0.090
ISI & PE	r= .725** p= 0.008	./.	./.
VI & PE	R= -.201 P= 0.530	./.	./.
Abbreviations: *correlation is significant at a 0.05 level, ** correlation is significant at the 0.01 level			

Summary of the perceptual evaluation of the participants' rhythm and the correlation between the evaluation and the rhythm measures

The perceptual evaluation of the participants' speech rhythm revealed significant differences between the groups for the Poem, Reading Passage and Spontaneous Speech. The ataxic dysarthric speakers' speech was perceived as having an abnormal rhythm compared to the control group.

Significant correlations between the perceptual evaluation and the results of the rhythm measures could be found for the nPVI, VarcoV and ISI.

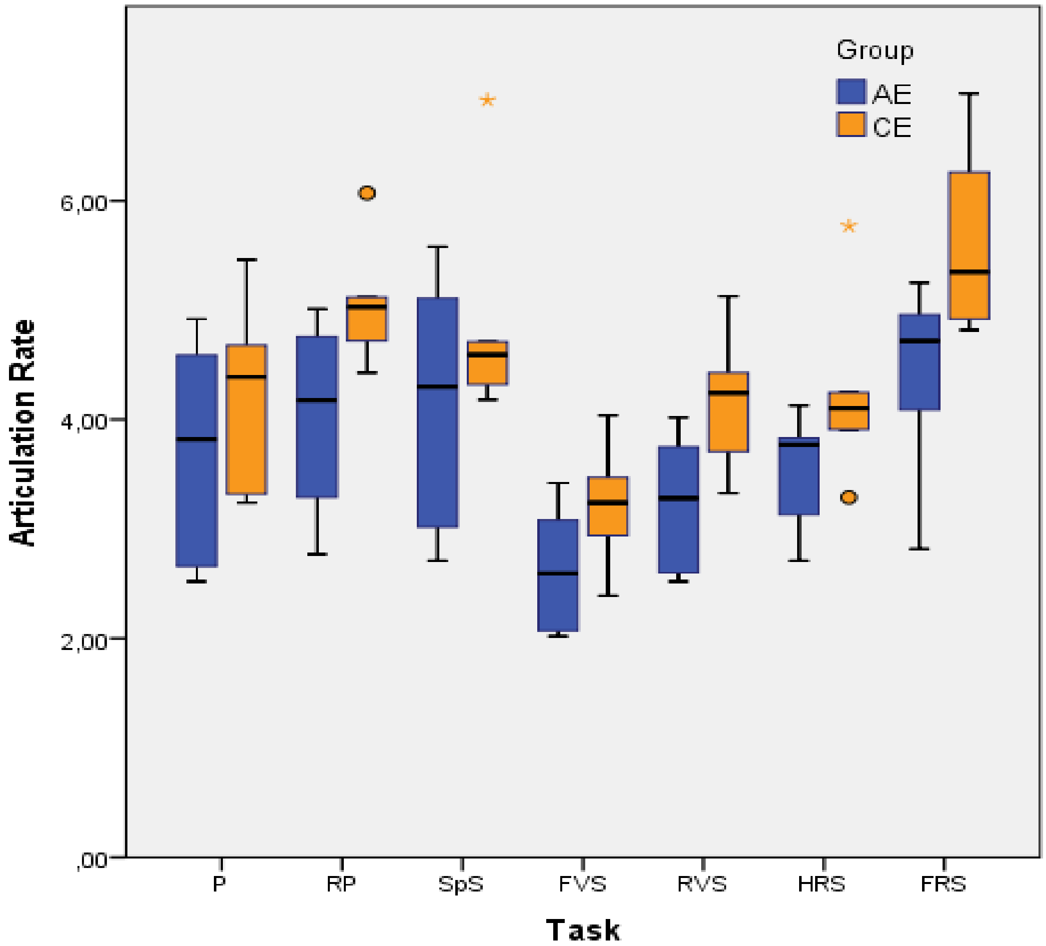
3.4 Articulation rate

The articulation rate per speaker for each task was calculated and correlated with the results for the rhythm metrics. Since speakers with ataxic dysarthria often present with a much lower articulation rate, compared to healthy control speakers, it is of importance that the rhythm metrics are not related to speech rate.

Figure 3.5 shows the means and range of the articulation rate for all tasks[9]. Articulation rate for the control group is generally higher, however a significant difference only occured for the Reduced Vowel Sentences (p = 0,037*, U = -2,08).

[9] The individual articulation rate for each speaker (syllables/ second), as well as the group mean and standard deviation can be found in the Appendix F.

Figure 3.5 Means and range for Articulation Rate for all tasks



Abbreviations: AE= English speaking ataxic dysarthric group, CE= English speaking control group, P = Poem, RP = Reading Passage, SpS = Spontaneous Speech Sample, FVS = Full Vowel Sentences, RVS = Reduced Vowel Sentences, HRS = Habitual Vowel Sentences, FRS = Fast Rate Sentences

Results show several significant correlations between articulation rate the rhythm metrics. For the ataxic dysarthric group the VarcoV and nPVI showed correlations for the Poem and Reading Passage, and additionally for the Full Vowel Sentences in the case of the nPVI. The control group showed significant correlations for the VarcoV and %V for the Reduced Vowel Sentences.

Table 3.10 Correlation of the articulation rate and the rhythm measures for the ataxic dysarthric group

	P	RP	SpS	FVS	RVS	HRS	FRS
nPVI	r= .886* p= 0.019	r= .886* p = 0.019	R= .657 p= 0.156	r= .886* p= 0.019	r= .371 p= 0.468	r= .600 p= 0.285	r= .600 p= 0.285
VarcoV	r= .886* p= 0.019	r= .829* p= 0.042	R= .543 p= 0.266	r= .543 p= 0.266	r= .714 p= 0.111	r= -.600 p= 0.285	r= .700 p= 0.188
%V	r= .257 p =0.623	r= .200 p= 0.704	R= .143 p= 0.787	r= -.143 p= 0.787	r= -.371 p= 0.468	r= -.600 p= 0.285	r= -.800 p= 0.188
ISI	r= -.543 p= 0.266	./.	./.	./.	./.	./.	./.
VI	r= .667 p= 0.148	./.	./.	./.	./.	./.	./.

Abbreviations: *correlation is significant at a 0.05 level, ** correlation is significant at the 0.01 level , P = Poem, RP = Reading Passage, SpS = Spontaneous Speech, FVS = Full Vowel Sentences, RVS = Reduced Vowel Sentences, HRS = Habitual Rate Sentences, FRS = Fast Rate Sentences

Table 3.11 Correlation of the articulation rate and the rhythm measures for the control group

	P	RP	SpS	FVS	RVS	HRS	FRS
nPVI	r= -.371 p= 0.468	r= -.086 p = 0.872	R= -.200 p= 0.704	r= -.314 p= 0.544	r= .143 p= 0.787	r= .257 p= 0.623	r= .227 p= 0.0502
VarcoV	r= .771 p= 0.072	r= -.429 p= 0.397	R= -.371 p= 0.468	r= .543 p= 0.266	r= .943** p= 0.005	r= .257 p= 0.623	r= .212 p= 0.556
%V	r= .486 p =0.329	r= .086 p= 0.872	R= .029 p= 0.957	r= -.029 p= 0.957	r= -.886* p= 0.019	r= -.714 p= 0.111	r= -.427 p= 0.190
ISI	r= .143 p= 0.787	./.	./.	./.	./.	./.	./.
VI	r= -.486 p= 0.329	./.	./.	./.	./.	./.	./.

Abbreviations: *correlation is significant at a 0.05 level, ** correlation is significant at the 0.01 level , P = Poem, RP = Reading Passage, SpS = Spontaneous Speech, FVS = Full Vowel Sentences, RVS = Reduced Vowel Sentences, HRS = Habitual Rate Sentences, FRS = Fast Rate Sentences

3.5 Summary of the results

Chapter three reported the results for the question which metrics would be most suited to highlight rhythmic differences in ataxic dysarthric speech compared to healthy control speakers and which speech task most sensitive to capture the rhythmic changes in ataxic dysarthric speech. Additionally, the correlation between severity of rhythmic disturbance as indicated by the different acoustic measures and the perceptual evaluation of speakers was investigated.

Results showed that some of the rhythm measures were able to show differences between the two speaker groups. The ISI and VarcoV showed significant group differences for the Poem. The VarcoV and nPVI additionally showed significant group differences for the Spontaneous Speech task. The other metrics did not show any significant group difference.

A general trend for more stress-timed values for the control group could be seen for most rhythm metrics; except the %V and VI. However, a high degree of overlap of the values between the groups and wide range of values was visible for the VI.

The three vowel based measures (nPVI, VarcoV, %V) were applied to several different tasks and not only the Poem. Some significant task differences for groups tasks were found in all three measures. However, the nPVI and VarcoV showed more significant task differences (nPVI: for the ataxic dysarthric group for the Poem and Reading Passage and Poem and Spontaneous Speech Sample, for the control group between the Reading Passage and Spontaneous Speech Sample, as well as between Full and Reduced Vowel Sentences, VarcoV: for the ataxic dysarthric group between the Poem and Reading Passage, the Reading Passage and Spontanteous Speech Sample, Full and Reduced Vowel Sentences, for the control group between the Poem and Reading Passage, Reading Passage and Spontaneous Speech, the Full and Reduced Vowel Sentences, as well as the Habitual and Fast Rate Sentences), whereas the %V only showed one (for the ataxic dysarthric group between the Reading Passage and Spontaneous Speech sample).

Three metrics presented significant correlations with the perceptual evaluation of the participants' speech rhythm (nPVI, VarcoV, ISI). The participants' speech rhythm was furthermore rated as being significantly different for the two groups for the Poem, Reading Passage and Spontaneous Speech tasks.

The VarcoV, nPVI and %V metrics, furthermore showed significant correlations with articulation rate (ataxic group for the Poem and Reading Passage (VarcoV, nPVI), Full Vowel Sentences (nPVI), the control group for the Reduced Vowel Sentences (VarcoV and %V)).

4. Discussion and Conclusion

The purpose of this study was to determine which methodologies are best suited to highlight rhythmic disturbances in ataxic dysarthria.

Specifically, it was investigated

1) which rhythm metrics are most suited to highlight rhythmic disturbances in ataxic speech compared to healthy control speech;
2) which speech tasks are most sensitive to capture rhythmic changes in ataxic dysarthric speech;
3) and which rhythm metrics best reflect the perceived impression of the speech disturbance of the ataxic dysarthric speakers.

The following discussion will address each of these questions in turn, starting with an evaluation of the relationship between rhythm metrics and articulation rate in order to be able to focus the remainder on those metrics that appear most valid to use in the investigation of ataxic dysarthria speech. The discussion will then turn to the group and task differences observed in this study and finally the correlation between the perceptual ratings and the rhythm metrics. The chapter will close with a conclusion, including implications and limitations of the study.

4.1 Discussion

4.1.1 Correlation of rhythm metric and articulation rate results

As discussed in earlier, metrics that correlate with articulation rate are not fit for purpose, because speakers with ataxic dysarthria often present with reduced rate compared to control speakers (Kent et al. 1997, Cannito & Marquardt 1997, Duffy 2012, Brown et al. 1970). In this study the VarcoV, nPVI and %V showed significant correlations (ataxic group for the Poem and Reading Passage (VarcoV, nPVI), Full Vowel Sentences (nPVI), the control group for the Reduced Vowel Sentences (VarcoV and %V)), suggesting that the metrics might not be robust against articulation rate. Due to the lack of previous studies investigating metrics in relation to different speaking tasks, no conclusions can be drawn as to whether this is suggestive of a general pattern or whether the results are idiosyncratic to the current participant group. The literature is inconclusive to that regard. White & Mattys (2007a, b), for example, found the VarcoV to be robust to articulation rate variation and in fact the most discriminating metric for rhythmically distinct languages. In contrast Liss et al. (2009) report high correlations for the VarcoV and articulation rate for their group of predominantly disordered speakers. They therefore eliminated the VarcoV from their final analysis, even though it showed to be an important metric in their first stepwise analysis. Researchers (White & Mattys 2007 a, b, Liss et al. 2009) found no correlations with articulation rate for the nPVI, however several studies found correlations between articulation rate and the %V metric (Dellwo & Wagner 2003, Wagner & Dellwo 2004, Liss et al. 2009).

It therefore appears that further studies are warranted to investigate the relationship between rhythm metrics and articulation rate in more detail across different tasks, with a particular focus on whether this relationship alters in disordered speakers.

For the moment, no clear decision can be made whether to exclude the VarcoV, %V and nPVI from the set of metrics or not, given the variable results reported in the literature and observed in this study. However, it appears sensible in the current discussion as well as for future studies to only consider results from tasks where the correlation was not significant.

4.1.2 Differences between the ataxic dysarthric and control group

Findings show, that the rhythm metric results for the control group for the nPVI, VarcoV and %V are in agreement with those published on English speakers in the literature (Grabe & Low 2002, Mok & Dellwo 2008, Liss et al. 2009, Low et al. 2000; Ramus et al. 1999, Spencelayh 2001, White & Mattys 2007a). The ISI values also agree with those previously published (Hartelius et al. 2000, Schalling & Hartelius 2004). The VI values for the Deterding (2001) study are lower compared to those found in the current study.

Differences pertaining to %V, nPVI, VarcoV and ISI occurred with regard to the performances of the speakers with dysarthria. With regard to %V Liss et al. (2009) found a higher %V value for their ataxic dysarthric group for sentence reading and phrases than the current study. Values for the ataxic dysarthric group for the current study varied across the tasks, however in an overall comparison across all tasks they showed a lower but comparable value to Liss et al. (2009). The VI has not been applied to ataxic dysarthric speech, therefore no comparison can be made.

The current results show three of the metrics (nPVI, VarcoV and ISI) were able to capture rhythmic differences between the ataxic dysarthric and control group for the Spontaneous Speech (nPVI, VarcoV, ISI) sample and the Poem (VarcoV), whereas the %V and VI did not show any significant difference. The %V only approached significance for the Spontaneous Speech task.

The nPVI has been successfully used for the differentiation of normal from ataxic speech (Liss et al. 2009, Stuntebeck 2002, Rosen et al. 2003) and it has been shown to be a useful tool in "capturing the rhythmic properties of ataxic speech" (Stuntebeck 2002:22). Whilst Liss et al. (2009) successfully used the nPVI for the discrimination of different dysarthria types, Kim, Kent & Weismer (2011) found opposite results. However, the authors (Kim et al. 2011) hypothesise that the use of the non normalized version of the PVI may have been partly responsible for the results, as compared to the Liss et al. (2009) study. Lowit (2014) could not ascertain group differences between dysarthric and control speakers with the nPVI, VarcoV and %V. However, she elucidated that her results may have been due to intra-group variability, sample size or the nature of the task (Lowit 2014).

In contrast to the results of Lowit (2014), the findings of the present study for the VarcoV are supported by the results of Liss et al. (2009). They found that the VarcoV metric can be successfully used for the differentiation between ataxic and normal speech. However, the VarcoV was eliminated from their final study, because of a correlation with articulation rate.

The results for the present study regarding the %V are in alignement with Lowit (2014), i.e. they did not reveal group differences. Contrary to that, the %V measure reliably identified group differences in the Liss et al. (2009) study. However, the %V was also eliminated from the final study of Liss et al. (2009), because of a correlation with articulation rate. This may be due to different procedual approaches of the studies (difference in speaking task, measures and speakers). As Wiget et al. (2010) stated the "problem in interpreting and comparing [...] results is that absolute rhythm scores can vary widely for the same languages between studies" (Wiget et al. 2010: 1560).

Knight (2011) found the %V to be a valid and reliable metric. She found it to be able to differentiate speaker languages and not influenced by speaker idiosyncrasities or time. Additionally, White & Mattys (2007 a, b) found the %V metric as most discriminative between rhythmically different languages. In contrast to that Arvaniti (2012) found inconsistent results across all tested metrics and mostly no significant task differences. She tested six different rhythm metrics (ΔC, %V, rPVI, nPVI, VarcoC and VarcoV) on spontaneous speech, sentences and story reading on six different languages (English, German, Greek, Italian, Korean and Spanish). Furthermore, she detected high inter-speaker variation and task dependency. The different results of the current study and Arvaniti (2012), migth partly be due to the fact, that Arvaniti (2012) carried out a cross-linguistic study, whereas the current study compared ataxic dysarthric and control speech.

As mentioned above, ataxic dysarthric speech is often described as having scanning speech or equally stressed syllables. It would therefore appear to have more equalised syllable durations than Standard English (Hartelius et al. 2000, Kent et al. 1979). Thus, a significant difference between the two groups for the Reduced Vowel Sentences was expected. Contrary to these expectations the nPVI, VarcoV and %V results for the two groups showed highly overlapping values with their means hardly differing. This might be due to the structured nature of the task, i.e. the sentences of these tasks were read out by the speakers, they were relatively short and had pauses in between. This allowed more time for planning which might have helped the ataxic dysarthric speakers to compensate for their rhythmic problems in this task as opposed to spontaneous speech.

Additionally to the group differences the general trend of the group values are worthy of discussion. Whereas the VarcoV and nPVI showed similar results with the ataxic dysarthric group showing more syllable-timed values as compared to the control group, the %V did not. Contrary to expectations the mean %V values for both groups were close together, and the ataxic group scored lower values for some tasks compared to the control group. As mentioned before, the ataxic dysarthric speech is often described as having equally stressed syllable durations (Hartelius et al. 2000, Kent et al. 1979), involving less reduction in vowel or consonant length, or a general reduction of vowels or consonants in syllables compared to healthy control speakers. Therefore, higher %V values would have been expected for the ataxic dysarthric group. The mild severity of the ataxic dysarthric participants could be mentioned as a reason for the lack of significant difference, as well as the different speech material used. Whereas other studies only applied the %V to read material the current study also used a Spontaneous Speech sample.

The control group showed a general trend of higher nPVI values representing a greater variability in vowel length compared to the ataxic dysarthric speakers. A generally higher nPVI for the control group as opposed to the ataxic group was also found by Stuntebeck (2002) for sentence reading. Additionally, Liss et al. (2009) also found higher nPVI values for their control speakers compared to the ataxic dysarthric speakers, with measured values being highly similar to the present results. They furthermore found in a first analysis of all their tested metrics significant group differences for the nPVI Hence, the nPVI is a useful tool for the differentiation between ataxic dysarthric and control speech in the current study.

For the ISI the coefficient of variation (CV) across inter-stress intervals was measured. The groups differed significantly from each other with the control group showing a much lower value ISI value. Based on the assumption that the speech of healthy control speakers has a tendency for isochrony, lower ISI CV values were expected for the control group than for the ataxic dysarthric group. For the speakers with dysarthria a higher value was found, which was also confirmed by the literature. For instance, Schalling & Hartelius (2004) found a notably higher variability for their Swedish speakers with ataxic dysarthria than for their controls. They reported on problems of their ataxic dysarthric speakers to keep regular stress beats, and concluded that a higher variability of ISI durations is one of the primary characteristics of ataxic dysarthria (cf. Schalling & Hartelius 2004:379).

Deterding (2001) found significant differences between his speaker groups and found the VI to be a suitable measure for the differentiation in rhythm along a stress- and syllable-timed scale. However, it was a cross-linguistic study with differing speech tasks. The VI has not been applied to ataxic dysarthric data and therefore no comparison can be made.

4.1.3 Differences between the tasks per group

In addition to group differences, the study also yielded differences in relation to elicitation task. Whereas the %V only showed one task difference for the ataxic group (Reading Passage / Spontaneous Speech), the nPVI and VarcoV showed several differences across both groups. However, they differed for the groups. The ataxic dysarthric group showed significant differences for the nPVI (Poem/ Reading Passage, Poem/ Spontaneous Speech, Full &/ Reduced Vowel Sentences), the VarcoV (Poem/ Reading Passage, Reading Passage/ Spontaneous Speech, Full &/ Reduced Vowel Sentences).

The control group revealed significant task differences for the nPVI (Reading Passage / Spontaneous Speech, Full & / Reduced Vowel Sentences) and the VarcoV (Poem / Reading Passage, Reading Passage / Spontaneous Speech, Full & / Reduced Vowel Sentences, Habitual & / Fast Rate Sentences) and none for the %V metric.

These differences have to be considered in the light of the correlations with articulation rate. The Spontaneous Speech task is the only task of the above, which did show a correlation with rate and can therefore only be considered for future research. Nevertheless, all results will be discussed.

Significant differences between the tasks may mirror the different requirements the tasks asks from the speaker (Kent et al. 1997). For the ataxic dysarthric group the three rhythm metrics did not completely result in significant differences for the same tasks. Research in dysarthria (Kent et al. 1997, Kent & Kent 2000, Lowit-Leuschel & Docherty 2001) found differences in performance across tasks for dysarthric and control speakers. The %V was not able to capture these differences as well as the other metrics. These showed differences between the tasks within the groups. This may suggest that the %V measure is not as sensitive to performance differences in speaking tasks as the other metrics.

The nPVI values for the control speakers for the Full Vowel Sentences and Reduced Vowel Sentences were slightly lower compared to the values reported in Low et al. (2000). However, Reduced Vowel Sentences showed significantly higher values compared to the Full Vowel Sentences for both speaker groups, similar to the results of Low et al. (2000). The VarcoV has not yet been used on Full Vowel Sentences and Reduced Vowel Sentences but results are comparable to the nPVI results.

Only one significant task difference was observed for the %V metric for the ataxic dysarthric group between the Reading Passage and Spontaneous Speech.

Contrary to expectations, the Reading Passage showed a significantly lower %V, indicating that the amount of vocalic intervals during Spontaneous Speech was higher than for the Reading passage.

Material of the current study was only controlled for speech rate in the Habitual Rate Sentences and Fast Rate Sentences. These sentences did not show significant task differences for either group for all rhythm metrics. Only the results for the VarcoV for the control group showed significance. However, the VarcoV also showed a significant correlation to articulation rate with the articulation rate, albeit for different tasks (Poem, Reduced Vowel Sentences). This suggests, that the VarcoV metric is sensitive to changes in tempo and not robust against rate (see also Liss et al. 2009).

Arvaniti (2009, 2012) found that the choice of speech tasks can significantly change the results of rhythm metrics, e.g. speech material which is more stress-timed would elicit more stress-timed rhythm scores, as well as more syllable-timed material would lead to more syllable-timed scores. The current study also used different speech tasks, which ranged from structured to unstructed. Interestingly, differences between the tasks occured mostly between read material and Spontaneous Speech for all metrics. This can be interpreted as a general trend, even though results have to be taken carefully, because of the correlations of the rhythm metrics and articulation rate. Differences between the Spontaneous Speech sample and other tasks, e.g. the Poem might be explained with requirements that the tasks ask from the speakers. As Kent et al. (1997) noted "different speaking tasks may emphasize different features of speech difficulties in persons with cerebellar disease" (Kent et al. 1997:76).

The Poem consists of five short and highly rhythmic structured sentences, it requires the speakers to adhere to a predetermined rhythm. Even though reading might be easier than producing Spontaneous Speech, the strict rhythm might have caused extra effort. For the Spontaneous Speech sample, the speakers were free in the production of speech, whether they used long or short phrases, vowels etc.. Spontaneous Speech requires a high amount of coordination of all speech processes from the speakers, together with the cognitive requirements of thinking about what to say. So even though the two tasks are quite different from each other, they show similar results for the rhythm metrics. This could be due to the high planning and controlling demands of the tasks.

The comparison of the results for the e.g. Poem and the Reduced Vowel Sentences show that the choice of speech task has an influence on the outcome. On the one hand reading might aid speakers with ataxic dysarthria to produce a speech rhythm comparable to control speakers (similar values for both groups in the Reduced Vowel Sentences for the %V, VarcoV and nPVI). On the other hand it seems to depend on the specific demands of the tasks, as adhering to a strict rhythm lead to a great difference between groups (differences in Poem). Therefore, the choice of speaking task in ataxic dysarthric research has to consider the specific demands the task makes, and tasks need to be chosen carefully in relation to the purpose of the assessment.

4.1.4 Correlation of rhythm metrics and perceptual evaluation

In order to answer the question which rhythm metric best reflects the perceived impression of the speech disturbance of the ataxic dysarthric speakers, a correlation between the results of the rhythm metrics and perceptual evaluation was carried out.

Results showed significant correlations for perceptual evaluation with the nPVI (Poem, Spontaneous Speech), VarcoV (Poem, Spontaneous Speech, Reading Passage) and ISI (Poem), however not for the other two metrics.

This indicates that speakers with lower nPVI, VarcoV or higher ISI values were perceptually rated as having a more abnormal speech rhythm. This is comparable with the findings of Stuntebeck (2002) for the nPVI. However, this does not align with the results of Lowit (2014). She could not find a correlation between the perceptual and acoustic measures nPVI, VarcoV, %V. Results for the latter metric are, however, comparable to the current results, as the %V did not show any significant correlation with the results of the perceptual evaluation of the participants' speech rhythm.

Hartelius et al. (2000) conclude that the absence of even inter-stress intervals (regular stress beats) adds to the perceptual impression of scanning speech. A similar picture was evident in the current study, with a significant correlation between the ISI CV values and the results of the perceptual evaluation of the participants´ speech.

Whilst results for the present study indicate that metrics can mirror the perceptual impression of the participants' speech, other studies reported differently (Arvaniti 2009, Lowit 2014).

4.2 Conclusion

This final chapter discussed the findings of the current study, which focused on rhythm in ataxic dysarthria. Different rhythm metrics were applied to a range of speech tasks and correlated with the perceptual evaluation of the participants' speech task. Results showed differences between the speaker groups for the rhythm metrics and tasks, as well as some correlations with the perceptual evaluation.

Results of the current study need to be interpreted carefully, because of the correlations with articulation rate. Further studies need to ensure that the relationship between the articulation rate and rhythm metrics are investigated in more details across different tasks, with a particular focus on whether this relationship alters in disordered speakers.

With this in mind the following conclusions are made: The first question of this study was which rhythm metric would be most suited to highlight rhythmic disturbances in ataxic speech compared to healthy control speech. Out of the five tested rhythm metrics three showed the best results: the VarcoV, nPVI, and ISI. The three metrics best highlighted the rhythmic differences between the two groups. The three metrics furthermore correlated with the perceptual evaluation of the participants' speech rhythm.

Even though the ISI showed group differences and the correlation with the perceptual evaluation, its use is limited. In order to determine inter-stress intervals, knowledge about the stress location within utterances is generally needed. It therefore, can only be applied to a small subset of tasks and not spontaneous speech.

The suitability of the nPVI and VarcoV becomes also evident with the answer to the question what the correlation is between severity of rhythmic disturbance as indicated by the different acoustic measures and the perceptual evaluation of speakers: The nPVI and VarcoV showed high correlations with the perceptual evaluation of the participants' speech rhythm.

The other metrics (%V, VI) did not show comparable results to the other rhythm metrics. They both failed to distinguish between the speaker groups and furthermore did not show any correlation with the perceptual impression of the participants' speech rhythm.

With regard to tasks, it might be considered, that different tasks may highlight different degrees of speech disturbance and researchers assume that the character of a certain dysarthria may vary with the speaking task (Brown & Docherty 1995, Kent et al. 1997, Kent & Kent 2000). Current results confirm task dependent performances for both speaker groups, as the groups showed significant differences between tasks. The nPVI and VarcoV showed significant task differences for both groups and therefore seems to be sensitive to capture task dependent differences.

As for the question which speech tasks are most sensitive to capture rhythmic changes in ataxic dysarthric speech, the Spontaneous Speech sample showed best results. The Spontaneous Speech sample was the one task showing most group differences for the various metrics, as well as a correlation with the percpeptual evaluation of the participants' speech rhythm. Furthermore, no metric showed correlations with articulation rate for this task. There are certain advantages to this task: Spontaneous Speech samples are the most representative of natural speech and with a given topic the comparability can be quite high.

Overall, the results of this study suggest that for further research into the rhythmic changes in ataxic dysarthria the nPVI and VarcoV are the most suitable metric, alongside a perceptual evaluation of the metric and Spontaneous Speech sample a suitable task,

4.3 Limitations and Implications

The findings of this study are clearly limited, as it included relatively few speakers who varied in the severity of their dysarthria and therefore it has to be seen if results can be generalised. Further studies into rhythm in ataxic dysarthria would benefit from a greater participant number and speakers with a higher severity of dysarthria.

On the other hand, the current study was encouraging to the point that is showed a good correlation between acoustic metrics and perceptual impression of the speakers. In addition, it has contributed further to the validation of these measures for research purposes. In addition, it has provided some guidance as to which metrics might be most differentiating, pointed to spontaneous speech as the most appropriate elicitation task and highlighted the fact that checks on the correlation between metric and articulation rate are essential before results can be reliably interpreted.

All of the above points should now be investigated in larger studies including more participants and a wider range of disorders. In addition, the successful differentiation of speaker groups with the rhythm metrics, as well as the quantification of rhythmic disturbance in an objective and reliable way, may be a useful tool for differential diagnosis and measuring therapeutic outcomes. It thus appears worthwhile to continue work to make acoustic rhythm analysis clinically applicable through automisation of the analysis.

REFERENCES

Abercrombie, D. (Ed.) (1965). *Studies in phonetics and linguistics*. London: Oxford University Press.

Abercrombie, D. (Ed.) (1967). *Elements of General Phonetics*. Edinburgh: Edinburgh University Press.

Ackermann, H. & Hertrich, I. (1994). Speech rate and Rhythm in Cerebellar Dysarthria: An Acoustic Analysis of Syllabic Timing. *Folia Phoniatrica et logopaedica*, *46*, 70-78.

Ackermann, H. & Hertrich, I. (2000). The contribution of the cerebellum to speech processing. *Journal of neurolinguistics*, *13*, 95-116.

Amerman J.D. & Parnell, M.M. (1992). Speech timing strategies in elderly adults. *Journal of Phonetics*, *20*, 65–76.

Anderson, A.H., Bader, M., Gurman Bard, E., Bayle, E., Doherty, G., Garrod, S., Isard, S., Kowtko, J., McAllister, J., Miller, J., Sotillo, C. & Thomson, H.S. (1991). The HCRC map task corpus. *Language and Speech*, *34*, 351-366.

Arvaniti, A. (2009). Rhythm, timing and the timing of rhythm. *Phonetica*, *66*, 46-63.

Arvaniti, A. (2012). The usefulness of metrics in the quantification of speech rhythm. *Journal of Phonetics*, *40*, 351-373.

Barry, W.J., Andreeva, B., Russo, M., Dimitrova, S. & Kostadinova, T. (2003). Do Rhythm Measures Tell us Anything about Language Type? *Proceedings of the 15th International Congress of Phonetic Sciences,* (pp. 2693-2696). Barcelona, Spain.

Barry, W.J. & Russo, M. (2003). Measuring rhythm. Is it separable from speech rate? In A. Mettouchi & G. Ferre (eds.) *Proceedings of the Acoustique Aquisition Iinterpretation Workshop "Prosodic Interfaces"*, (pp. 15-20). Nantes, France.

Boersma, P. & Weenink, D. (2013). Praat: doing phonetics by computer [Computer program]. Version 5.1.25, retrieved from http://www.praat.org/

Boutsen, F., Bakker, K. & Duffy, J. (1997). Subgroups in ataxic dysarthria. *Journal of medical speech-language pathology*, *5* (1), 27-36.

Brown, A. & Docherty, G.J. (1995). Phonetic variation in dysarthric speech as a function of sampling task. *European Journal of Disorders of Communication*, *30* (1), 17-36**.**

Brown, J.R., Darley, F. L. & Aronson, A.E. (1970). Ataxic dysarthria. *International journal of neurology*, *7*, 302-318**.**

Cannito, M.P. & Marquardt, T.P. (1997). Ataxic Dysarthria. In M.R. McNeil (ed.), *Clinical Management of Sensori-motor Speech Disorders,* (pp. 217-249). New York: Thieme.

Cauldwell, R. (2002). The functional irrhythmicality of spontaneous speech: A discourse view of speech rhythms. *Apples - Journal of applied languages studies*, *2* (1), 1-24.

Chenery, H.J., Ingram, J.C.L. & Murdoch, B.E. (1990). Perceptual analysis of speech in ataxic dysarthria. *Australian journal of human communications disorders*, *18* (1), 19-18.

Clark, J. & Yallop, C. (1995). Prosody. In J. Clark & C. Yallop, (eds.), *An Introduction to Phonetics and Phonology*, (pp.328-363). Oxford: Basil Blackwell.

Crystal, D. & Davy, D. (1969) *Investigating English Style*. Hong Kong: Longman.

Darley, F.L., Aronson, A.E. & Brown, J.R. (1969a). Clusters of deviant speech dimensions in the dysarthrias. *Journal of Speech and Hearing Research*, *12*, 462-469.

Darley, F. L., Aronson, A.E. & Brown, J.R. (1969b). Differential diagnostic patterns of dysarthria. *Journal of Speech and Hearing Research, 12*, 246-269.

Darley, F.L., Aronson, A.E. & Brown, J.R. (eds.) (1975). *Motor Speech Disorders*. Philadelphia: W.B. Saunders Company.

Dauer, R.M. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics*, *11*, 51-62.

Dauer, R.M. (1987). Phoneitc and phonological components of language rhythm. *Proceedings of the XIth International Congress of Phonetic Sciences,* (pp. 447-450). Talin, Estonia.

Dellwo, V. (2004). Rhythm and Speech Rate: A variation coefficient for C. In P. Karnowski & I. Szigeti (eds.), *Proceedings of the 38s linguistic Colloquium,* (pp.231-241). Piliscsaba, Hungary.

Dellwo, V. (2006). Rhythm and Speech Rate: A variation coefficient for delta C. In: P. Karnowski & I. Szigeti (eds.), *Language and language-processing,* (pp. 231-241). Frankfurt am Main, Peter Lang,.

Dellwo, V. (2008a). The role of speech rate in perceiving speech rhythm. *Proceedings at the Speech Prosody,* (pp. 375-378). Campinas, Brazil.

Dellwo, V. (2008b). Influences of language typical speech rate on the perception of speech rhythm. *Journal of the Acoustical Society of America, 123* (5), 3427.

Dellwo, V. & Wagner, P. (2003). Relations between language rhythm and speech rate. *Proceedings of the 15th International Congress of Phonetic Sciences,* (pp. 471-474). Barcelona, Spain.

Deterding, D. (1994). The intonation of Singapore English. *The Journal of the International Phonetic Association*, *24* (2), 61-72.

Deterding, D. (2001). Letter to the Editor: The measurement of rhythm: a comparison of Singapore and British English. *Journal of Phonetics*, *29*, 217-230.

Dilley, L., Wallace, J. & Heffner, C. (2012). Perceptual isochrony and fluency in speech by normal talkers under varying task demands. In O. Niebuhr & H. Pfitzinger (eds.), *Prosodies: Context, Function, and Communication, Language, Context, and Cognition series,* (pp. 237-258). Berlin: Walter deGruyter.

Duffy, J.R. (1995). Ataxic Dysarthria. In J.R. Duffy (ed.), *Motor Speech Disorders* (pp. 145-165). St. Louis: Elsevier.

Duffy, J.R. (2012). Ataxic Dysarthria. In J.R. Duffy (ed.), *Motor Speech Disorders*. St. Louis: Elsevier.

Durr, A. (2010). Autosomal dominant cerebellar ataxias: polyglutamine expansions and beyond. *Lancet Neuol*, *9*, 885-894.

Eriksson, A. (1991). *Aspects of Swedish Speech Rhythm.* Unpublished Gothenburg Monographs in Linguistic 9, University of Gothenburg, Gothenburg, Sweden.

Evidente, V.G.H. , Gwinn- Hardy, K.A., Caviness, J.N. & Gilman S. (2000). Hereditary Ataxias. *Mayo Clinic Proceedings, 75*, 475-490.

Fant, G., Kruckenberg, A. & Nord, L. (1989). Rhythmical Structures in Text Reading. A Language Contrasting Study. *Eurospeech, 89*, 498-501.

Fant, G., Kruckenberg, A. & Nord, L. (1991). Durational correlates of stress in Swedish, French and English. *Journal of Phonetics, 19*, 351-365.

Fonteyn, E.M.R., Schmitz-Hübsch, T., Verstappen, C.C., Baliko, L., Bloem, B.R., Boesch, S., Bunn, L., Charles, P., Dürr, A., Filla, A., Giunti, P., Globas, C., Klockgether, T., Melegh, B., Pandolfo, M., De Rosa, A., Schöls, L., Timmann, D., Munneke, M., Kremer, B.P.H. & van de Warrenburg, B.P.C. (2010). Falls in Spinocerebellar Ataxias: Results of the EuroSCA Fall Study. *Cerebellum*, *9*, 232-239.

George, D. & Mallery, P. (2002). SPSS for Windows Step by Step: A Simple Guide and Reference, 11.0 Update (4th ed.). Boston: Allyn & Bacon.

Gentil, M. (1990). Dysarthria in Friedreich Disease. *Brain and Language*, *38*, 438-448.

Ghez, C. (1991). The Control of Movement. In E.R. Kandel, J.H. Schwartz & T.M. Jessel (eds.), *Principles of neural science* (3rd ed., pp. 553-547). New York: Elsevier.

Grabe, E. (30.03.2010). *MS Excel Worksheet for the calculation of PVI values,* retrieved from: http://www.phon.ox.ac.uk/files/people/grabe

Grabe, E. & Low E.L. (2002). Durational variability in speech and the rhythm class hypothesis. In N. Warner & C. Gussenhoven (eds.), *Proceedings of the Laboratory Phonology, 7,* (pp.515-546) Berlin, Germany: Mouton de Gruyter.

Handel, S. (1989). Chapter 11 Rhythm. In S. Handel (ed.), *Listening - An Introduction to the Perception of Auditory Events,* (pp. 383-460), Cambridge: MIT Press.

Hartelius, L., Runmaker, B., Andersen, O., & Nord L. (2000). Temporal Speech Characteristics of Individuals with Multiple Sclerosis and Ataxic Dysarthria: ‚Scanning Speech' Revisited. *Folia Phoniatrica et Logopaedica, 52*, 228-238.

Hartelius, L., Theodoros, D., Cahill, L., Lillvik M. (2003). Comparability of Perceptual Analysis of Speech Characteristics in Australian and Swedish Speakers with Multiple Sclerosis. *Folia Phoniatrica et logopaedica, 55*, 177-188.

Hertrich, I. & Ackermann, H. (1993). Dysarthria in Friedreich's ataxia: syllable intensity and fundamental frequency patterns. *Clinical Linguistics & Phonetics, 7*(3), 177-190**.**

Kent, R. (2000). Research on Speech Motor Control and Its Disorders: A Review And Prospective. *Journal of Communication Disorders, 33*, 391-428.

Kent, R. & Netsell, R. (1975). A case study of an ataxic dysarthric: Cineradiographic and spectrographic observations. *Journal of Speech and Hearing Disorders, 40* (1), 115-134**.**

Kent, R., Netsell, R. & Abbs, J.H. (1979). Acoustic characteristics of dysarthria associated with cerecellar disease. J*ournal of Speech and Hearing Disorders, 22*, 627-648.

Kent, R. & Rosenbek, J.C. (1982). Prosodic disturbance and neurologic lesion. *Brain and Language, 15*, 259-291.

Kent, R., Weismer, G., Kent, J.F. & Rosenbek, JC. (1989). Toward phonetic intelligibility testing in dysarthria. *Joural of Speech and Hearing Disorders, 54*(4), 482-499.

Kent, R., Kent, J.F., Rosenbek, J.C., Vorperian, H. K. & Weismer, G. (1997). A speaking task analysis of the dysathria in cerebellar disease. *Folia Phoniatrica et Logopaedica, 49*, 63-82.

Kent, R., Kent, J.F., Duffy, J.R., Weismer, G. (1998). The dysarthrias: Speech - voice profiles, related dysfunctions, and neuropathology. *Journal of Medical Speech-Language Pathology, 6*(4), 165-211.

Kent, R., Kent, J.F, Duffy, J.R., Thomas, J.E., Weismer, G. & Stuntebeck, S. (2000). Ataxic dysarthria. *Journal of Speech, Language, and Hearing Research*, *43*, 1275-1289.

Kim, Y., Kent, R., & Weismer, G. (2011). An acoustic study of the relationships among neurologic disease, dysarthria type, and severity of dysarthria. *Journa of Speech, Language and Hearing Research, 54,*417-429.

Klockgehter T,. Bürk K., Auburger G., & Dichgans, J. (1995). Klassifikation und Diagnostik der degenerativen Ataxien. *Nervenarzt*, *66*, 571-581.

Knight, R.-A. & Cocks, N. (2007). Rhythm in the speech of a person with right hemisphere damage: Applying the pairwise variability index. *Advances in Speech–Language Pathology, 9* (3), 256-264.

Knight, R.-A. (2011). Assessing the temporal reliability of rhythm metrics. *Journal of the International Phonetic Association*, *41* (3), 271-281.

Kohler, K. J. (1995). Articulatory Reduction in Different Speaking Styles. *Proceedings of the 13th International Congress of Phonetic Sciences,* (Vol. 2, pp. 12-19). Stockholm, Sweden.

Ladefoged, P. (1967). *Three Areas of of Experimental Phonetics.* London: Oxoford University Press.

Leuschel, A. & Docherty, J.G. (1997). Prosodic Assessment of Dysarthria. In D.A. Robin, K.M. Yorkston & D.R. Beukelman. (eds), *Disorders of Motor Speech - Assessment, Treatment and Clinical Characterization,* (pp. 155-180). Baltimore: Paul Brookes Publishers.

Lindblom, B. (2010). Fo curves – smooth, seamless yet pulsed? *Proceedings from Fonetik 2010*, June 2-4, 51 - 56, Lund, Sweden.

Linebaugh, W.C. & Wolfe, E.V. (1984). Relationships between Articulation Rate, Intelligibility, and Naturalness in Spastic and Ataxic Speakers. In M.R. McNeil, J.C. Rosenbek & A.E. Aronson (eds.), *The Dysarthrias - Physiology, Acoustics, Perception, Management,* (pp. 197-205). San Diego: College Hill Press.

Liss, J.M., Spitzer, S.M., Caviness, J.N., Adler, C. & Edwards, B.W. (2000). Lexical boundary error analysis and ataxic dysarthria. *Journal of the Acoustic Society of America, 107* (6), 3415-3424.

Liss, J.M., White, L., Mattys, S.L-, Lansford, K., Lotto, A.J., Spitzer, S.M. & Caviness, J.N. (2009). Quantifying Speech Rhythm Abnormalities in the Dysarthrias. J*ournal of Speech, Language, and Hearing Research***,** *52*, 1334-1352.

Love, R.J. (1995). Motor Speech Disorder. In H.S. Kirshner (ed.), *Handbook of Neurological Speech and Language Disorder,* (pp. 41-56). New York: Marcel Dekker Inc..

Low, E.L. & Grabe, E. (1995). Prosodic patterns in Singapore English, *Proceedings of the 13th International Congress of Phonetic Sciences,* (pp. 636-639). Stockholm, Sweden.

Low, E.L. (1998). *Prosodic prominence in Singapore English.* Doctoral dissertation, University of Cambridge, England.

Low, E.L., Grabe, E. & Nolan, F. (2000). Quantitative Characterizations of Speech Rhythm: Syllable-Timing in Singapore English. *Language and Speech, 43* (4), 377-401.

Lowit, A. (2014). Quantification of rhythm problems in disordered speech: a re-evaluation. *Philosophical Transactions of the Royal Society B,* 369, 1-13

Lowit, A., Miller, N., Poedjianto, N. & McCall, J. (2001). A comparison of traditional and modified DDK tasks with connected speech. In B. Maasen (ed.) *Speech motor control in normal and disordered speech: 4th International Speech Motor Conference,* (pp. 204-207). Nijmegen, the Netherlands.

Lowit, A., Miller, N. & Poedjianto, N. (2003). Characteristics Of Performance Change In Dysarthria: Clinical Perspectives. *Journal of Clinical Speech and Language Studies, 12/13*, 87-107.

Lowit-Leuschel, A. (1997). *Prosodic Impairment in Dysarthria: An Acoustic Phonetic Study.* Doctoral Dissertation. University of Newcastle upon Tyne, Newcastle upon Tyne, England.

Lowit -Leuschel, A. & Docherty, G.J. (2001). Prosodic variation across sampling tasks in normal and dysarthric speakers. *Logopedics, Phoniatrics & Vocology, 26,* 151-164.

Lowit, A., Kuschmann, A. & MacLeod, J.M. (2010). Sentence Stress in Ataxic Dysarthria — A Perceptual and Acoustic Study. *Journal of Medical Speech-Language Pathology, 18* (4), 77–82.

Ludlow, C.L. & Bassich, C.J. (1983). The Result of Acoustic and Perceptual Assessment of Two Types of Dysarhtria. In W.R. Berry (ed.), *Clinical Dysarthria,* (pp. 121-154). San Diego: College Hill Press.

Mok, P. & Dellwo, V. (2008). Comparing native and non-native speech rhythm using acoustic rhythmic measures: Cantonese, Beijing Mandarin and English. *Proceedings of the Speech Prosody Campinas,* (pp.423-426) Campinas, Brazil.

Murry, T. (1983a). The Production of Stress in Three Types of Dysarthric Speech. In W.R. Berry (ed.), *Clinical Dysarthria,* (pp. 69-83), San Diego: College Hill Press.

Murry, T. (1983b). Treatment of Ataxic Dysathria. In W.H. Perkins (ed.), *Dysathria and Apraxia - Current Therapy of Communication Disorder,* (pp.79-89), Stuttgart: Georg Thieme Verlag.

Mundt, J.C., Snyder, P.J., Cannizzaro, M.S., Chappie, K. & Geralts, D.S. (2007). Voice acoustic measures of depression severity and treatment response collected via interactive voice response (IVR) technology. *Journal of Neurolinguistics*, *20*, 50-64.

Nestell, R. (1984). A Neurobiologic View of the Dysarthrias. In: M.R. McNeil, J.C. Rosenbek & A.E. Aronson (eds.), *The Dysarthrias - Physiology, Acoustics, Perception, Management* (pp. 1-36), San Diego: College Hill Press.

Nishio, M. & Niimi, S. (2001). Speaking rate and its components in dysarthric speakers. *Clinical Linguistics & Phonetics, 15* (4), 309-317.

Opal, P. & Zoghbi, H.Y. (2002). The hereditary ataxias. In A.K. Asbury, G.M. McKhann, W.I. McDonald, P.J. Goadsby & J.C. McArthur (eds.), *Diseases of the nervous system - Clinical Neuroscience and Therapeutic Principles,* (3rd ed., pp. 1880-1895), Cambridge: Cambridge University Press.

Patterson, David (2000): *A linguistic approach to pitch range modelling.* Doctoral Thesis, University of Edinburgh, Scotland.

Peterson, G.E. & Lehiste, I. (1960). Duration of syllable nuclei in English. *Journal of the Acoustical Society of America, 32,* 693-703.

Petrov, A. & Anderson, J. (2005). The dynamics of scaling: A memory-based anchor model of category rating and absolute identification. *Psychological Review, 112* (2), 383-416.

Pike, K. (1945). General Characteristics of Intonation. In K. Pike (ed.), *The intonation of American English,* (pp. 20-43), Ann Arbor, University of Michigan Press.

Pompino-Marschall, B. (ed.) (1990). *Die Silbenprosodie: ein elementarer Aspekt der Wahrnehmung von Sprachrhythmus und Sprechtempo.* Tübingen: Niemeyer, de Gruyter.

Ramus, F., Nespor, M. & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition, 73,* 265-292.

Ramus, F. (2002). Acoustic Correlates of Linguistic Rhythm: Perspectives Speech Prosody. In B. Bel & I. Marlin (eds.), *Proceedings of Speech Prosody, Laboratoire Parole et Langage,* (pp.115-120), Aix-en-Provence, France.

Roach, P. (1982). On the distinction between 'stress-timed' and 'syllable-timed' languages. *Linguistic Controversies*. 73-79.

Robertson, S.J. and Thomson, F. (1986). *Working with Dysarthrics, A Practical Guide to Therapy for Dysarthria*, Oxford: Winslow Press.

Rosen, K.M., Kent, R. & Duffy, J.R. (2003). Lognormal distribution of pause length in ataxic dysarthria. *Clinical linguistics & phonetics, 17* (6), 469–486.

Rosen, K.M, Folker, J.E., Murdoch, B.E., Vogel, A.P., Cahill, L.M., Delatycki, M.B. & Corben, L.A. (2011). Spectral measures of the effects of Friedreich's ataxia on speech. *International Journal of Speech-Language Pathology*, *13* (4), 329-334.

Schalling, E. & Hartelius, L. (2004). Acoustic analysis of speech tasks performed by three individuals with spinocerebellar ataxia. *Folia Phoniatrica et logopaedica, 56*, 367-380.

Schalling, E., Hammarberg, B. & Hartelius, L. (2008). A Longitudinal Study of Dysarthria in Spinocerebellar Ataxia (SCA): Aspects of Articulation, Prosody, and Voice. *Journal of Medical Speech - Language Pathology*, *16*,103-117.

Scherer, K. R. & Ekman, P. (eds.) (1982). *Handbook of Methods in Nonverbal Behaviour Research*. Cambridge: Cambridge University Press.

Schmitz-Hübsch, T., Coudert, M., Bauer, P., Giunti, P., Globas, C., Baliko, L., Filla, A., Mariotti, C., Rakowicz, M., Charles, P., Ribai, P., Szymanski, S., Infante, J., van de Warrenburg, B., Dürr, A., Timmann, D., Boesch, S., Fancellu, R., Rola, R., Depondt, C., Schöls, L., Zdienicka, E., Kang, J.-S., Döhlinger, S., Kremer, B., Stephenson, D.A., Melegh, D.B., Pandolfo, M., di Donato, S., Tezenas du Montcel, S., Klockgether, T. (2008). Spinocerebellar ataxia types 1, 2, 3, and 6. Disease severity and nonataxia symptoms. *Neurology*, *71*, 982-989.

Schöls, L., Bauer, P., Schmidt, T., Schulte, T. & Riess, O. (2004). Autosomal dominant cerebellar ataxias: clinical features, genetics, and pathogenesis. *Lancet Neurol*, *3*, 291- 304.

Setter, J. (2000). *Rhythm and Timing in Hong Kong English*. Doctoral Thesis. University of Reading, Reading, England.

Setter, J., Stojanovik, V., Van Ewijk, L., & M. Moreland (2007). Affective prosody in children with Williams syndrome. *Clinical Linguistics & Phonetics,* 21 (9): 659-672.

Smith, L. & Rafiqzad, K. (1979). English for cross-cultural communication: the question of intelligibility. *Tesol Quarterly, 13*(3), 371-380.

Spencelayh, B. (2001). *Comparing rhythmic variation in four British dialects.* Unpublished Undergraduate Project. University of York, England.

Spencer, K.A. & Rogers, M.A. (2005). Speech motor programming in hypokinetic and ataxic dysarthria. *Brain and Language, 94,* 347–366.

Stuntebeck, S. (2002). *Acoustic ananlysis of the prosodic properties of ataxic speech.* Unpublished master's thesis, University of Wisconsin, Madison, USA.

Wagner, P. & Dellwo, V. (2004). Introducing YARD and Re-Introducing isochrony to rhythm research. *Proceedings of Speech Prosody*, Nara, Japan.

Walker, H.K. (1990). The Cerebellum, In H.K. Walker, W.D. Hall & J.W. Hurst, *Clinical Methods: The History, Physical, and Laboratory Examinations* (pp. 356-359), Boston: Butterworths.

Weismer, G. (1984). Acoustic Description of Dysarthric Speech. *Seminars in Speech an Language, 5* (4), 293-313.

Wiget, L., White, L., Schuppler, B., Grenon, I., Rauch, O. & Mattys, S. (2010). How stable are acoustic metrics of contrastive speech rhythm? *Acoustical Society of America, 127* (3), 1559-1569.

White, L., & Mattys, S. (2007a). Calibrating rhythm: First language and second language studies. *Journal of Phonetics, 35,* 501–522.

White, L., & Mattys, S. (2007b). Rhythmic Typology and variation in the first and second languages. In P. Prieto, J. Mascarœ & M.-J. Solé (eds.), *Segmental and prosodic issues in Romance phonology, Current Issues in Linguistic Theory series* (pp. 237–257), Amsterdam: John Benjamins.

White, Karessa (2012). *The acoustic characteristic of Ataxic Dysarthria*, Unpublished Master Thesis, University of Florida, USA.

Yorkston, K.M., Beukelmann, D.R., Minifie, F.D. & Sapir, S. (1984). Assessment of stress Patterning. In: M.R. McNeil, J.C. Rosenbek & A.E. Aronson (eds.), *The dysarthrias - physiology, acoustics, perception, management* (pp. 129-162), San Diego: College Hill Press.

Zellner-Keller, B. & Keller, E. (in press). The chaotic nature of speech rhythm. In Ph. Delcloque & V.M. Holland (eds.), *Integrating Speech Technology in Language Learning,* Swets & Zeitlinger.

Ziegler, W. & Wessel, K. (1996). Speech timing in ataxic disorder: Sentence production and rapid repetitive articulation. *Neurology, 47*, 208-214**.**

Ziegler, W., Vogel, M., Groene, B., Schröter-Morasch H. (1998). *Dysarthrie - Grundlagen, Diagnostik, Therapie*. Stuttgart: Thieme.

Ziegler, Wolfram. (2002). Task-Related Factors in Oral Motor control: Speech and Oral Diadochokinesis in Dysarthria and Apraxia of Speech. *Brain and Language*, *80*, 556-575.

Ziegler, W. (2003). Review. Speech motor control is task-specific: Evidence from dysarthria and apraxia of speech. *Aphasiology*, *17*(1), 3-36.

APPENDIX

A. Overview on which participant took part in which task

Speaker	P	RP	SpS	FRS	HRS	RVS	FVS	PE
AE_1	X	X	X	X	X	X	X	X
AE_2	X	X	X	X	X	X	X	X
AE_3	X	X	X	X	X	X	X	X
AE_4	X	X	X	X	X	X	X	X
AE_5	X	X	X	-	-	X	X	X
AE_6	X	X	X	X	X	X	X	X
CE_1	X	X	X	X	X	X	X	X
CE_2	X	X	X	X	X	X	X	X
CE_3	X	X	X	X	X	X	X	X
CE_4	X	X	X	X	X	X	X	X
CE_5	X	X	X	X	X	X	X	X
CE_6	X	X	X	X	X	X	X	X

Abbreviation: AE= English speaking ataxic dysarthric group, CE= English speaking control group, P = Poem, RP = Reading Passage, SpS = Spontaneous Speech, FVS = Full Vowel Sentences, RVS = Reduced Vowel Sentences, HRS = Habitual Rate Sentences, FRS = Fast Rate Sentences, PE = Perceptual Evaluation

B. nPVI values for all speakers for all tasks

high value indicates stress-timing, low value syllable timing

Speaker	P	RP	SpS	FRS	HRS	RVS	FVS
AE_1	50,04	45,89	47,75	46,19	45,32	52,5	37,7
AE_2	70,42	61,3	66,5	57,01	58,47	61,79	47,34
AE_3	62,71	58,64	50,43	56,96	59,65	70,26	45,92
AE_4	62,73	54,71	47,07	67,92	58,29	73,18	39,84
AE_5	52,79	45,46	47,82	-	-	62,05	33,16
AE_6	56,39	53,34	43,19	50,9	42,5	59,56	38,6
CE_1	54,82	58,95	69,82	57,09	62,67	52,5	37,7
CE_2	67,26	49,21	54,63	51,71	45,27	51,53	41,69
CE_3	69,86	52,61	72,82	61	51,54	62,64	44,69
CE_4	63,4	62,12	78,14	61,81	58,52	70,33	51,16
CE_5	74,87	59,61	55,87	60,05	54,24	69,84	47,35
CE_6	80,75	53,06	61,84	59,93	49,29	77,41	47,91
Mean AE	59,18	53,22	50,46	55,80	52,85	63,22	40,43
SD AE	7,53	6,49	8,20	8,16	8,23	7,49	5,33
Mean CE	68,49	55,93	65,52	58,60	53,59	64,04	45,08
SD CE	9,02	5,01	9,55	3,73	6,31	10,43	4,82

Abbreviation: AE= English speaking ataxic dysarthric group, CE= English speaking control group, P = Poem, RP = Reading Passage, SpS = Spontaneous Speech, FVS = Full Vowel Sentences, RVS = Reduced Vowel Sentences, HRS = Habitual Rate Sentences, FRS = Fast Rate Sentences

C. VarcoV values for all speakers for all tasks

high value indicates stress-timing, low value syllable timing

Speaker	P	RP	SpS	FRS	HRS	RVS	FVS
AE_1	49,03	40,02	45,61	43,58	42,62	42,55	34,51
AE_2	61,14	52,15	54,90	56,10	57,86	54,41	39,61
AE_3	62,87	54,91	67,24	56,21	65,41	56,87	38,33
AE_4	54,73	51,88	54,58	57,60	56,68	54,90	34,23
AE_5	41,04	41,55	40,49	-	-	46,52	30,04
AE_6	48,93	42,27	63,40	52,22	39,22	53,54	32,41
CE_1	63,78	51,63	64,07	63,58	55,19	58,92	37,00
CE_2	57,51	48,84	57,63	50,48	40,87	50,32	29,77
CE_3	63,63	47,09	86,79	49,10	45,14	51,84	33,33
CE_4	73,51	57,04	80,51	52,20	51,12	56,99	36,70
CE_5	62,11	54,17	85,85	51,61	46,97	49,53	34,77
CE_6	76,26	46,50	65,17	51,51	43,35	51,20	32,64
Mean AE	52,96	47,13	54,37	53,14	52,36	51,47	40,43
SD AE	8,27	6,54	10,17	5,71	11,03	5,62	3,59
Mean CE	66,13	50,88	76,59	53,08	47,11	53,13	45,08
SD CE	7,20	4,17	12,55	5,26	5,26	3,86	2,72

Abbreviation: AE= English speaking ataxic dysarthric group, CE= English speaking control group, P = Poem, RP = Reading Passage, SpS = Spontaneous Speech, FVS = Full Vowel Sentences, RVS = Reduced Vowel Sentences, HRS = Habitual Rate Sentences, FRS = Fast Rate Sentences

D. %V values for all speakers for all tasks

High value indicate syllable timing, low values stress-timing,

Speaker	P	RP	SpS	FRS	HRS	RVS	FVS
AE_1	55,81	41,78	51,59	39,47	46,13	44,25	48,53
AE_2	49,2	38,12	41,7	38,48	34,07	38,61	44,41
AE_3	45,96	42,55	53,57	16,67	33,91	40,13	42,49
AE_4	51,58	43,28	51,04	29,89	33,46	48,73	47,48
AE_5	24,66	36,54	52,51	-	-	46,77	40,48
AE_6	43,46	40,97	48,31	45	46,9	51,57	44,89
CE_1	49,15	42,5	41,52	32,42	25,82	35,92	48,53
CE_2	46,53	41,05	42,37	39,12	41,61	48,19	46,99
CE_3	43,9	37,17	51,45	37,53	22,45	20,18	32,28
CE_4	45,52	40,07	47,56	35,81	44,53	39,69	39,64
CE_5	42,1	44,33	45,06	39,4	42,73	43,30	40,57
CE_6	42,13	46,04	33,43	33,72	43,01	40,82	41,37
Mean AE	45,11	40,51	49,79	33,90	38,89	45,01	44,71
SD AE	10,91	2,56	4,34	11,05	6,97	5,01	3,01
Mean CE	44,89	41,86	43,56	36,34	36,69	38,02	41,56
SD CE	2,74	3,16	6,14	2,86	9,83	9,64	5,81

Abbreviation: AE= English speaking ataxic dysarthric group, CE= English speaking control group, P = Poem, RP = Reading Passage, SpS = Spontaneous Speech, FVS = Full Vowel Sentences, RVS = Reduced Vowel Sentences, HRS = Habitual Rate Sentences, FRS = Fast Rate Sentences

E. Values for the ISI & VI for all speakers for the Poem

For the ISI high values indicate syllable-timing, low values stress-timing

for the VI high value indicates stress-timing, low syllable-timing

Speaker	ISI	VI*100
AE_1	51,65	43,00
AE_2	36,77	81,00
AE_3	36,39	80,00
AE_4	38,46	58,00
AE_5	43,55	75,00
AE_6	27,71	75,00
CE_1	35,05	59,00
CE_2	34,59	84,00
CE_3	27,18	61,00
CE_4	30,64	64,00
CE_5	32,34	69,00
CE_6	32,53	80,00
Mean AE	39,09	68,60
SD AE	8,00	15,07
Mean CE	32,06	69,30
SD CE	2,88	10,20
Abbreviation: AE= English speaking ataxic dysarthric group, CE= English speaking control group		

F. Articulation rate for each speaker

for each task presented in syllables/ second

Speaker	P	RP	SpS	FRS	HRS	RVS	FVS
AE_1	2,66	2,77	2,71	2,82	2,71	2,60	2,02
AE_2	4,92	5,01	5,11	4,96	4,13	4,02	3,42
AE_3	4,59	4,73	5,58	5,25	3,77	3,34	2,83
AE_4	4,30	4,76	5,10	4,72	3,83	3,75	3,08
AE_5	2,52	3,29	3,50	-	-	2,52	2,07
AE_6	3,34	3,62	3,02	4,09	3,13	3,22	2,36
CE_1	5,46	6,07	6,92	6,98	5,77	5,13	4,04
CE_2	3,32	4,72	4,71	5,59	4,16	3,71	2,94
CE_3	4,23	5,05	4,68	6,26	4,25	4,43	3,47
CE_4	4,55	5,01	4,32	5,11	4,05	4,41	3,30
CE_5	3,24	4,43	4,51	4,92	3,29	3,33	2,39
CE_6	4,68	5,12	4,18	4,82	3,91	4,08	3,18
Mean AE	3,72	4,03	4,17	4,37	3,51	3,24	2,63
SD AE	1,02	0,93	1,24	0,97	0,58	0,60	0,57
Mean CE	4,25	5,07	4,89	5,61	4,24	4,18	3,22
SD CE	0,85	0,55	1,02	0,85	0,82	0,63	0,55

Abbreviation: AE= English speaking ataxic dysarthric group, CE= English speaking control group, P = Poem, RP = Reading Passage, SpS = Spontaneous Speech, FVS = Full Vowel Sentences, RVS = Reduced Vowel Sentences, HRS = Habitual Rate Sentences, FRS = Fast Rate Sentences

G. Instructions and rating form

for the listener judges for the perceptual evaluation of the participants speech rhythm for the Poem, Reading Passage and Spontaneous Speech Sample (one rating form per speech task)

Instruction:
You will hear a Limerick, an extract of a reading passage and some spontaneous conversation. Please rate the speaker's rhythm. The data should be rated on the 5 point scale with 1 indicating a normal and 5 a highly abnormal rhythm.
Rating form 4th year students

	1= no abnormality				5= highly abnormal
	1	**2**	**3**	**4**	**5**
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					

Rating from 2nd year students

Question: How disturbed is the rhythm?

How disturbed is the rhythm?

- ☐ 1. - No abnormality
- ☐ 2. -
- ☐ 3. -
- ☐ 4. -
- ☐ 5. - Highly abnormal

H. Poem: Individual rhythm ratings of the listener judges R1-R18 and mean perceptual evaluation value for the Poem

Rater / Speaker	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	Mean Rating
AE_01	2	3	2	4	4	3	4	4	2	2	3	3	2	4	3	3	4	3	3,06
AE_02	2	1	2	2	3	2	2	2	2	4	2	3	3	2	2	2	2		2,24
AE_03	2	1	1	2	2	1	2	1	1	1	2	1	2	2	1	1	1	1	1,39
AE_04	3	2	3	3	4	2	4	4	2	3	2	2	4	2	3	2	3	3	2,83
AE_05	4	4	4	4	4	4	4	4	3	4	2	4	5	3	3	3	3	3	3,61
AE_06	2	2	3	3	2	2	2	2	2	2	3	2	2	4	2	2	3	2	2,33
CE-01	3	2	3	4	3	2	4	1	2										2,00
CE_02	2	2	2	2	2	2	2	2	1										1,56
CE_03	2	2	2	2	2	2	2	2	2										1,11
CE_04	1	1	1	1	1	1	1	1	1										1,11
CE_05	1	1	1	1	1	1	1	1	1										1,11
CE_06	1	1	2	1	1	1	1	1	1										1
Mean AE																			2,59
SD AE																			0,90
Mean CE																			1,31
SD CE																			0,39
AE_02_2	3	1	2	4	2	1	2	2	2										2,11
CE_05_2	1	1	1	1	1	1	1	1	1										11

Abbreviation: AE= English speaking ataxic dysarthric group, CE= English speaking control group, 1 indicates normal, 5 highly abnormal rhythm, R1-R18 = rater 1 to 18

I. Reading Passage: Individual rhythm ratings of the listener judges R1-R9 and mean perceptual evaluation value for the Reading Passage

	R1	R2	R3	R4	R5	R6	R7	R8	R9	Mean Rating
AE_01	3	4	4	4	4	4	4	4	3	3.78
AE_02	3	2	3	3	2	3	2	2	1	2.33
AE_03	2	1	2	2	2	1	2	2	1	1.67
AE_04	3	3	3	3	3	3	4	3	3	3.11
AE_05	2	3	3	3	3	3	3	3	4	3.00
AE_06	3	2	3	4	3	4	3	3	2	3.00
CE-01	3	2	3	4	3	2	4	1	2	2.67
CE_02	2	2	2	2	2	2	2	2	1	1.89
CE_03	2	2	2	2	2	2	2	2	2	2.00
CE_04	1	1	1	1	1	1	1	1	1	1.00
CE_05	1	1	1	1	1	1	1	1	1	1.11
CE_06	1	1	2	1	1	1	1	1	1	1.11
Mean AE										2.81
SD AE										0.61
Mean CE										1.63
SD CE										0.67
AE_02_2	2	2	2	3	2	2	3	2	2	2.22
CE_05_2	1	1	1	1	1	1	2	1	1	1.11
Abbreviation: AE= English speaking ataxic dysarthric group, CE= English speaking control group, SD= Standard Deviation, 1 indicates normal, 5 highly abnormal rhythm, R1-R9 = rater 1 t 9										

J. Spontaneous Speech: Individual rhythm ratings of the listener judges R1-R9 and mean perceptual evaluation value for the Spontaneous Speech Sample

	R1	R2	R3	R4	R5	R6	R7	R8	R9	Mean Rating
AE_01	3	4	3	4	4	4	4	4	4	3,78
AE_02	1	1	1	1	2	1	2	3	1	1,11
AE_03	1	2	2	2	2	2	2	3	1	1,89
AE_04	2	2	2	3	3	3	4	4	3	2,89
AE_05	2	2	2	3	3	3	4	4	3	2,89
AE_06	2	2	2	3	3	2	4	4	2	2,67
CE-01	2	2	2	1	4	2	3	2	2	2,22
CE_02	1	1	1	1	1	1	1	2	1	1,11
CE_03	1	1	1	2	2	1	1	2	1	1,33
CE_04	1	1	1	1	1	1	1	2	1	1,11
CE_05	1	1	1	1	1	1	1	2	1	1,11
CE_06	2	1	1	2	2	1	2	2	1	1,56
Mean AE										2.54
SD AE										0.92
Mean CE										1.41
SD CE										0.40
AE_02_2	1	1	1	1	1	1	1	2	1	1,11
CE_05_2	1	1	1	1	1	1	1	1	1	1

Abbreviation: AE= English speaking ataxic dysarthric group, CE= English speaking control group, SD= Standard Deviation, 1 indicates normal, 5 highly abnormal rhythm, R1-R9 = rater 1 t 9

K. Instruction and rating form for the intelligibility jugdements

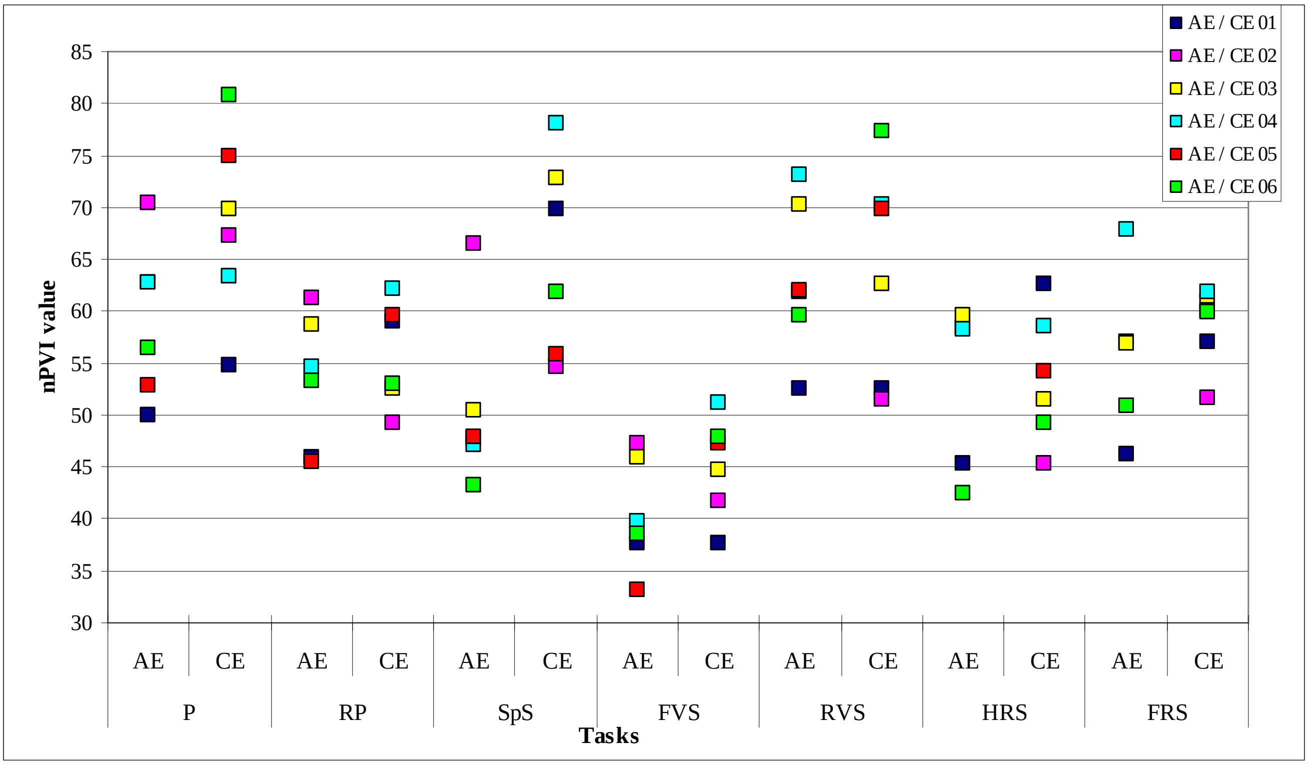
You will hear a Limerick. Please rate how the intelligibility the person is to you and give a value between 1 (not intelligible) and 100 (totally intelligible).

Question: How intelligible is the person? - give % value (1-100)

How intelligible is the person? - give % value (1-100)

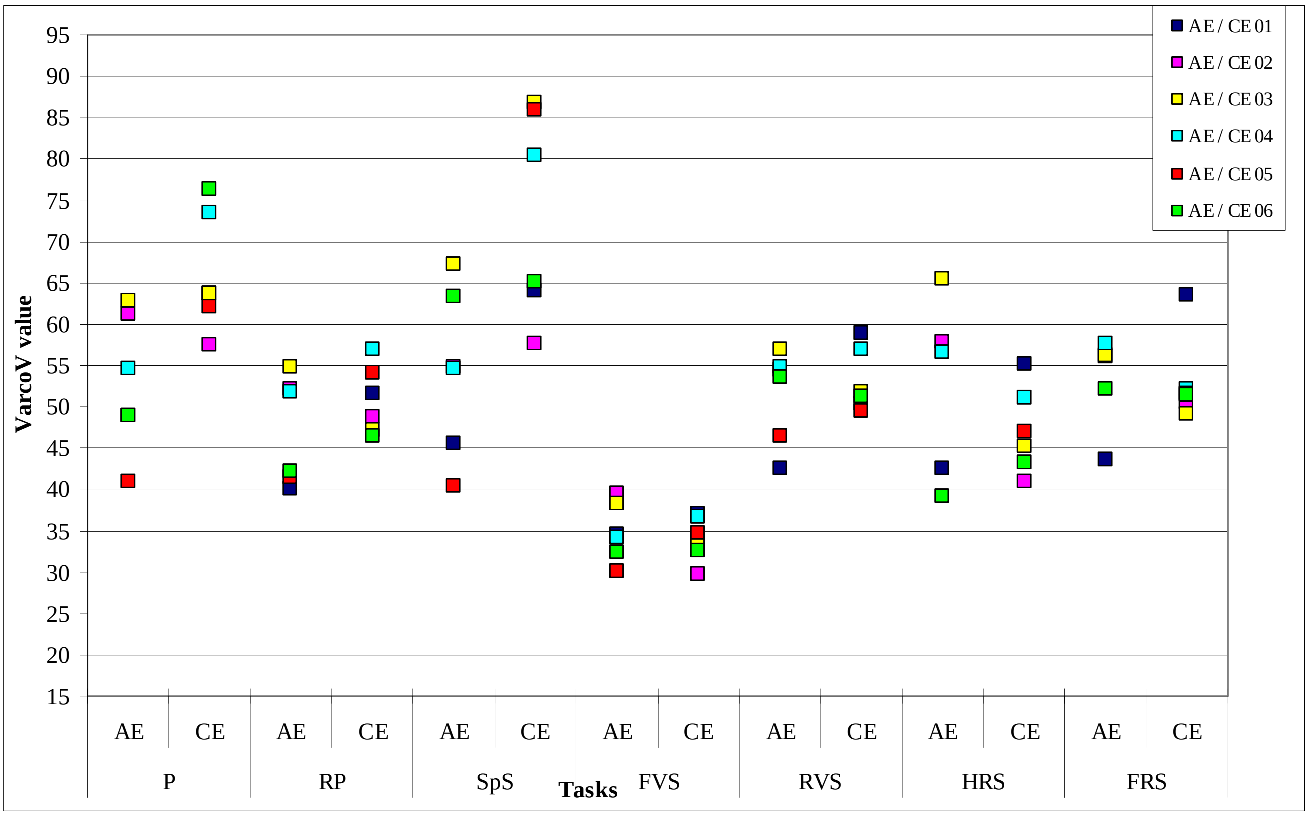
Answer:

L. Figure L Individual nPVI values for all tasks



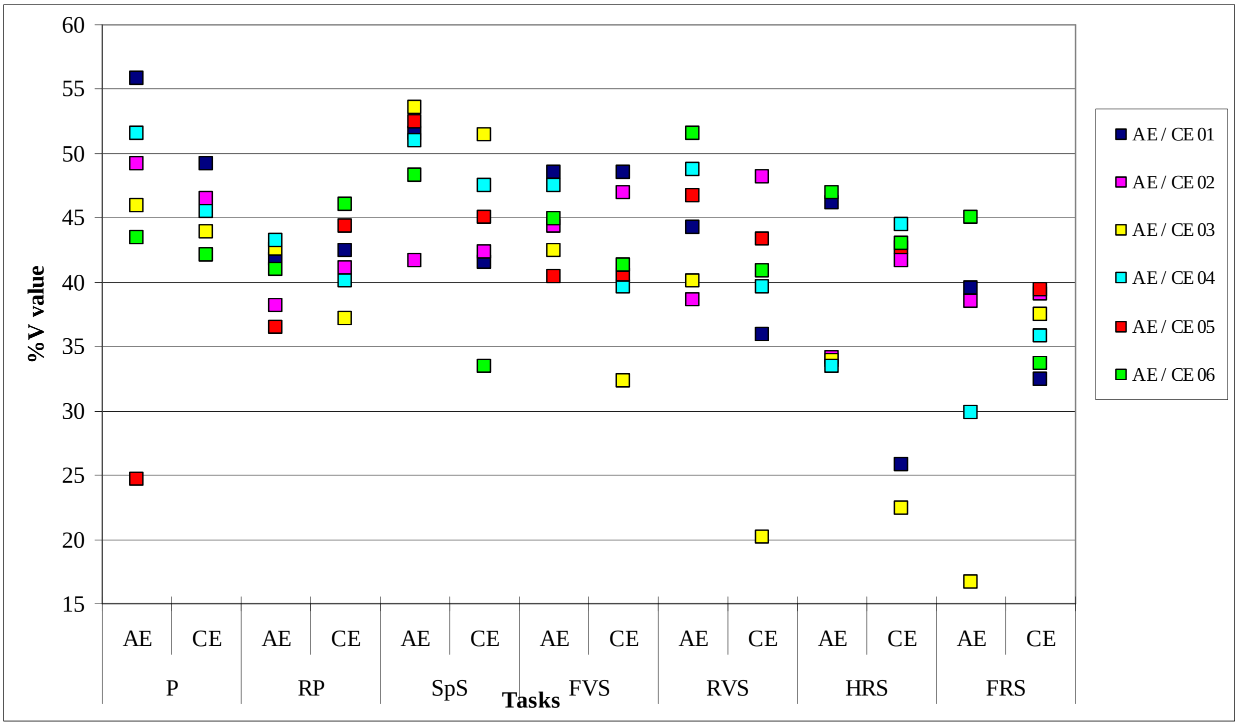
Abbreviations: AE= English speaking ataxic dysarthric group, CE= English speaking control group, P = Poem, RP= Reading Passage, SpS= Spontaneous Speech, FVS = Full Vowel Sentences, RVS = Reduced Vowel Sentences, HRS = Habitual Rate Sentences, FRS = Fast Rate Sentences

M. Figure M Individual VarcoV values for all tasks



Abbreviations: AE= English speaking ataxic dysarthric group, CE= English speaking control group, P = Poem, RP = Reading Passage, SpS = Spontaneous Speech, FVS = Full Vowel Sentences, RVS = Reduced Vowel Sentences, HRS = Habitual Rate Sentences, FRS = Fast Rate Sentences

N. Figure N Individual %V values for all tasks



Abbreviations: AE= English speaking ataxic dysarthric group, CE= English speaking control group, Full Vowel Sent. = Full Vowel Sentences, Reduced Vowel Sent. = Reduced Vowel Sentences, Normal Rate Sent. = Normal Rate Sentences, Fast Rate Sent. = Fast Rate Sentences