Development and Validation of a test for Assessment of Prosody in Children

A project funded by AIISH Research Fund (2018-2019)

Sanction No.: SH/CDN/ARF-SLP-4/2018-19 dated 9.10.18

Total Fund: Rs. 4,98,000/-

Project Report

Principal Investigator

Dr. N. Sreedevi

Professor & Head

Department of Clinical Services

All India Institute of Speech and Hearing

Mysuru-06

Co-Investigator

Dr.Sangeetha. M

Clinical Reader

Department of Clinical Services

All India Institute of Speech and Hearing

Mysuru-06

Research Officer

Ms.Priyadharshini.V.

All India Institute of Speech and Hearing

Manasagangothri, Mysuru 570006

Acknowledgements

Our sincere gratitude to our former Director, Prof. S. R Savithri and our present Director, Prof. M. Pushpavathi, All India Institute of Speech and Hearing, Mysuru, for sanctioning and providing all the support and help in completion of the project.

We extend our gratefulness to bio statisticians Dr. Vasanthalakshmi.M.S. and Mr. Santhosha. C.D. and for their help in statistical analysis. We also extend our thanks to Ms. Priyanka. T and Ms. Divyashree.K.N. for their participation as judges in data analysis. We thank the school authorities and children of Gangothri Public School and ACME School, Mysore, for their co-operation and participation during data collection. We thank all the participants, staff of DCS and other staff of the institute who helped in one way or the other in completion of this project report.

Dr. N. Sreedevi
Principal Investigator

Dr. Sangeetha. M
Co Investigator

Ms. Priyadharshini.V.
Research Officer

Table of Contents

LIST OF TABLES

List of Figures

CHAPTER I

INTRODUCTION

1.1 Introduction on suprasegmentals/prosody in speech

1.2 Different dimensions of prosody of speech

1.2.1 Stress

1.2.2 Rhythm

1.2.3 Intonation

1.3 Relation between language and prosodic skills

1.3.1 Prosody and language disorders

1.4 Assessment of prosody

1.5 Need for the Study

1.6 Aim

1.7 Objectives

1.1 Introduction on suprasegmentals/prosody in speech

Speech is a verbal means of communication and its components include articulation, voice, resonance and prosody (Owens, 2016). The suprasegmental information called as prosody, provides vital information to listeners in obtaining the meaning from spoken utterances including speech rate, volume, pitch and rhythm of speech (Ito, 2014). Speech perception involves simultaneous encoding of both segmental ns suprasegmental information in the input signal, which is significant for effective spoken communication (Abercrombie, 1967). As Pisoni and Sawusch (1975) suggest, "Prosody may serve as the interface between low level segmental information and higher levels of grammatical structures in speech". Real time physical events (such as changes in frequency, intensity and duration) in the acoustic speech signal serve as cues for the perception of prosodic features.

The four suprasegmental aspects are stress, intonation, juncture and rhythm. Stress refers to the accentuation or emphasis, laid on syllable or word. Intonation refers to variations in pitch over a time period. Juncture refers to the boundaries between the phonological units, signaled by segmental modifications. Rhythm refers to the pattern of accents / stress on a string of syllables. Rhythm gives shape/skeleton to a sentence, an idea of the length of a sentence and melody. It also marks the beginning and ending of a phrase and leads to the ease of pronunciation.

In the past, several investigators have attempted to explore the rhythmic features in various languages. Several researchers (Ainsworth, 1972; Lehiste, 1977; Martin, 1979; Nakatani, Connor & Aston, 1981; Mok & Dellwo, 2008; Fuchs, 2016) have studied rhythm in English. Indian languages which were studied include Tamil (Asher, 1985) and Kannada (Jayanthi Ray, 1993; Savithri, Jayaram, Kedarnath & Sanjay, 2006; Savithri, Maharani, Sanjay & Deepa, 2007; Savithri, Sreedevi & Kavya, 2009a; Savithri, Sreedevi & Kavya, 2009b;

Savithri, Sreedevi, Jayakumar & Kavya, 2010; Priyanka, 2018). Savithri, Maharani, Sanjay and Deepa (2007) investigated speech rhythm in 12 Indian languages, namely, Assamese, Bengali, Gujarati, Kashmiri, Marathi, Oriya, Punjabi, Rajasthani, Kodava, Malayalam, Tamil and Telugu.

1.2 Dimensions of Prosody

Children acquire linguistic skills as they frame the repertoire of speech sounds and lexical-syntactic components (Bowerman, 1982; Guasti, 2017). Perception of prosody is vital to understand spoken utterances, especially to appreciate intricacies of a language. Prosody is important during child language acquisition as it provides distinct cues for early communication. Prosodic emphasis or stress helps the listeners to focus attention towards specific segments (syllables/words/ phrases/sentences) which is cued by an increase in pitch, duration and amplitude (of the emphasized segments) (Breen, Kaswer, Van Dyke, Krivokapic & Landi, 2016). The major dimensions of prosody include stress, intonation and rhythm.

1.2.1 Stress in speech

Stress indicates the most prominent syllable/word in an utterance. Acoustically, the stressed syllable is characterized by increased fundamental frequency, duration, and intensity relative to the surrounding syllables (Ohala, 1977; Balusu, 2001; Collins & Mees, 2003; Keane, 2006). Stress can be defined in terms of effort involved in speech production, with the stressed syllables requiring more effort than unstressed syllables (Lehiste, 1970). During the developmental period stress perception and production were found to be near adult-like around 8 years of age, however reaches the target (adult-like) resemblance by adolescence (Ballard, Djaja, Arciuli, James & Doorn, 2012).

Studies have determined the acoustic correlates of stress in many languages, for example, duration was found to be the major acoustic correlate in Swedish (Westin, Buddenhagen & Obrecht, 1966), Estonian (Lehiste, 1968a) and Italian (Bertinetto, 1980). Fundamental frequency was the main acoustic correlate of stress in English (Morton & Jassem, 1965), French (Rigault, 1962) and Polish (Jassem, Morton & Steffen-Botog, 1968). In Indian languages, fundamental frequency and duration were found to be the major acoustic correlates of stress in Hindi (Agarwal, Ghosh & Savithri, 2007). In Dravidian languages, duration alone was reported to be the prime cue in Kannada (Savithri, 1987; Rajupratap, 1991; Savithri, 1999a), Tamil (Balasubramanian, 1981) and Konkani (Kumar & Bhat, 2009).

In terms of grammatical aspects, the content words are stressed more often than functional words (Kenworthy, 1987). Stress in an utterance is also reported to occur in certain positions in an utterance for example, the default phrasal stress in English is reported to occur on the last major word (Russell, 1997), where as in Kannada it reportedly occurs on the first few syllables (Manjula, 1997).

Listeners use stress cues to differentiate various meanings of words (that is, “PREsent” versus “preSENT”, the *former* refers to a *noun* meaning a gift and the *latter* functions as a *verb* meaning to deliver a message or act of gifting) and to identify word boundaries (Nakatani & Schaffer, 1978). Cues such as change in pitch and syllable lengthening prior to the boundary help listeners in comprehending phrase boundaries (de Pijper & Sanderman, 1994; Streeter, 1978; Choi, Haswegawa & Cole, 2005; Cumming, 2010). The perception and production of stress have been studied among the clinical population and marked deficits have been reported among individuals with apraxia (Gelfer, 1980), stuttering (Wingate, 1985), children with aphasia and autism (Baltaxe, 1984), learning disability (Highnam & Morris, 1987) and children with language impairments (John & Howarth, 1965; Nickerson, 1975).

1.2.2 Rhythm in Speech

Speech rhythm is broadly defined as recurring spectral–temporal patterns within the acoustic signal and it functions universally to reduce the computational load of recognizing spoken language (Borrie, Lansford & Barrett, 2017). Rhythm perception plays an important role in the process of lexical segmentation which is crucial for speech processing (Jusczyk & Luce, 2002). Several measures have been used to classify speech rhythm in languages. These include standard deviation and syllable duration (Abercrombie, 1964), inter-stress interval (Roach, 1982), % vocalic duration, ∆ vocalic and ∆ consonantal (Rams, Nesper, & Mechler, 1999). Of these % vocalic and ∆ consonantal could classify rhythm types in languages.

Low (1998) introduced Pairwise Variability Index (PVI) which measured the durational difference between successive vocalic intervals and successive intervocalic intervals. PVI could classify rhythm of languages which were traditionally classified as syllable timed, stress timed and mora-timed (Low, Grabe, & Nolan, 2000). High vocalic and intervocalic PVI's implied stress timed rhythm, low vocalic and intervocalic PVI's implied mora-timed and low vocalic and high intervocalic PVI's implied syllable timed rhythm. If a language has simple syllabic structure, for example, VC or CCV, the durational difference between the simplest and most complex syllable is not wide and is referred to as syllable-timed rhythm. Any syllable would be less than 330 ms in duration. If a language has complex syllables, for example, V and CCCVCC, the difference between syllables can be very large and is referred to as stress-timed rhythm. The difference in duration between syllables can be very wide ranging from 150 ms (in monosyllables) to 620 ms (in the word 'strength') in English (stress-timed language). If the syllable is still simpler, for example, VC or CV, then the durational difference between syllables is negligible and is referred to as mora-timed rhythm. In stress-timed languages, the time intervals between stressed segments are said to be near equal, where as in syllable-timed languages successive syllables are said to be of near-equal length.

In mora-timed languages successive morae* are said to be near equal in duration (* Morae are syllables consisting of one short vowel and any preceding onset consonant in the form of CV structure). Thus, mora-timed languages are more similar to syllable-timed languages than stress-timed languages (Savithri, 2007).

In the Indian context, Savithri, Jayaram, Kedarnath and Sanjay (2006) reported Kannada to be a mora-timed language based on PVI measures in adults. The speech rhythm in native Kannada speaking children was found to be unclassified based on PVI measures (Savithri, Sreedevi & Kavya, 2009a; Savithri, Sreedevi & Kavya, 2009b; Savithri, Sreedevi, Jayakumar & Kavya, 2010).In stress-based / stress-timed languages, like English or Dutch, listeners exploit metrical stress to guide their segmentation decisions (Mattys, White, & Melhorn, 2005; McQueen, Norris, & Cutler, 1994). In particular, the presence of strong syllables, those receiving relative stress through longer duration, fundamental frequency change, increased loudness, and a relatively full vowel, informs the onset of a new word (Cutler & Butterfield, 1992; Cutler & Norris, 1988). Mattys et al. (2005) opined that rhythm perception is the key for speech processing in adverse listening conditions. Rhythm may be broadly defined as the "structure of a sequence". (Allen 1975). The structure is described in terms of relationship among specific units (syllables/words/phrases/sentences) that constitute the structure.

If there exists no rhythm, speech breaks down leading to 'dysprosodia'. The sense of rhythm is not properly developed or is disrupted in the hearing impaired, stuttering, cluttering, dysarthria, aphasia, verbal apraxia. (Schlanger, Schlanger, & Gerstman, 1976; Parkhurst& Levitt, 1978; Starkweather, 1987). These clinical populations need to be evaluated and rehabilitated for which the clinicians needs to know about the development of rhythm. The developmental studies of rhythm, in the literature, indicate that the segmental timing shows a developmental trend in children. By 15 months of age, rhythm starts developing and continues till 12 years of age (Atkinson - King, 1973).

1.2.3 Intonation in Speech

Intonation is defined as the variations of pitch or fundamental frequency (F0) in speech as a function of time (Collier, 1991). Crystal (1973) suggested that intonation is an important mediator in interaction situations. Intonation conveys information specific to human emotion/ attitude. Every language has discrete and well-defined pitch patterns or intonation contours which signify the intended meaning of the speaker. Some intonation patterns are not specific in meaning and do not convey information about the implications of the speaker's attitude / feelings. Such patterns are referred to as "colourless/neutral/normal intonation" (Pike, 1945). On the other hand, few intonation contours contain a meaning with implied sense of speaker during the speech act. By 18 months of age, typically developing children begin to use adult-like intonation patterns in their utterances specific to the language environment (Menyuk, 1972).

Attempts have been made in the past to study the aspects of intonation in various languages. In this regard, studies by Crystal (1973) in English, Sandner (1981) in German, Thorsen (1978) in Danish and Williams (1985) in Welsh have the evidence that infants develop various intonation patterns over a period of time and they respond to suprasegmental properties at an early age at the expense of other segmental aspects/linguistic features. Literature on the deficits in perception and production of intonation patterns had suggested evident impairments in various communication disorders such as hearing impairment (Monsen, 1979; Sussman & Hernandez, 1979), developmental apraxia of speech (Tallman & Crary, 1993), cerebellar dysarthria and hypokinetic dysarthria (Darley, Aronson & Brown, 1969b).

1.3 Relation between language and prosodic skills

Accurate perception of prosody is crucial for successful language acquisition. In the recent past, some of the studies suggest the existence of close relation between music and language functions in the course of development. Koelsch (2011) came up with a hypothesis, which states that " the human brain, especially during early age, does not treat language and music as strictly separate domains, but it tends to treat language as a special case of music." Brandt, Slevc and Gebrian (2012) concluded parallel development of musical/prosodic and linguistic abilities based on similar electrophysiological markers of syntax processing in language and music. Music not only plays a major role in acquiring musical aspects of native culture, but in addition it also helps in language acquisition (Brandt et al., 2012).

There are a lot of similarities between language and music, namely the acoustic categories (phonemes and tones), which are organized in systematic sequences, with respect to certain regularities. The acquisition of such regularities involves similar learning mechanisms. Electrophysiological studies have indicated shared cognitive resources and also similar underlying neural substrates for the processing of semantics (Steinbeis & Koelsch, 2008), syntax (Jentschke & Koelsch, 2009) and prosody (Magne, Schön & Besson, 2006). Among the above, prosody is presumed to be the area having a relatively stronger overlap with language. Such links between prosody and language acquisition suggest that the difficulties with prosody perception are likely to be the cause for delay in language acquisition.

1.3.1 Prosody and Language Disorders

Prosody perception is strongly linked to language development in the literature, and is reported to be impaired in several clinical populations. As stated earlier, prosody is conveyed by a combination of different acoustic cues (pitch/ amplitude/ duration). Hence, a single source of difficulty with prosody perception likely leads to auditory processing deficits.

However, impairments in auditory processing can be present for one dimension in the presence of preserved processing in other dimensions. In particular, impaired pitch perception can co-occur with preserved duration perception and vice versa (Kidd, Watson & Gygi, 2007). Similarly, research on amusia (lack of music perception) has shown that highly impaired memory for pitch sequences can co-occur with preserved memory for durational sequences (Hyde & Peretz, 2004). A prosody perception deficit in a given individual, therefore, could reflect impaired pitch perception or duration perception or both.

Phonemic and prosodic awareness are independent predictors of word reading (Jimenez-Fernandez, Gutierrez-Palma, & Defior, 2015; Wade-Woolley & Heggie, 2015; Wade-Woolley, 2016), suggesting that prosody perception forms a separate dimension of linguistic skill relevant to reading acquisition. Dyslexia has been linked to impaired prosody perception (Goswami, Gerson, & Astruc, 2010; Mundy & Carroll, 2012; Wade-Woolley, 2016; Wood & Terrell, 1998), Literature suggests impairment in prosody skills among children with Specific Language Impairment (CwSLI) with parallel deficits in both prosody and language (Marshall, Brown, Ramus & Van der lely, 2009; Corriveau & Goswami, 2009; Clement, Plachou, Beland, Motte & Samson, 2015; Sallat & Jentschke, 2015; Mari, Scorpecci, Reali & Alatri, 2016; Akhila, 2016).

1.4 Assessment of prosody

Although there exists many widely available standardized tests of segmental speech perception for individuals of all ages (Nilsson, Soli& Sullivan, 1994; Wilson, 2003; Killion, Niquette, Gudmundsen, Revit & Banerjee, 2004) there are comparatively few assessment tools available for researchers and clinicians interested in testing suprasegmental aspects of speech. As a consequence, prosody perception research has been carried out using a wide variety of in-house methods developed within single laboratories, making comparison across studies difficult. These include perceptual matching tasks such as matching low-pass filtered sentences or indicating whether the prosodic structure of low-pass filtered sentences match unfiltered target sentences (Wood & Terrell, 1998; Fisher, Plante, Vance, Gerken & Glattke, 2007; Cumming, Wilson, Leong, Colling & Goswami, 2015). Participants have also been asked to match the stress pattern of a nonsense phrase like “DEEdee DEEdee” with a spoken target phrase like “Harry Potter” (Whalley & Hansen, 2006; Goswami, 2010; Holliman, Wood & Sheehy, 2012; Mundy & Carroll, 2012). These tests have the advantage of isolating the suprasegmental elements of speech. However, these tests are not publicly available.

Koike and Asp (1981) developed the Tennessee Test of Rhythm and Intonation Patterns (T-TRIP) to assess rhythm and intonation. Kalathottukaren, Purdy and Ballard (2015) reviewed the assessment tools that have been developed to assess prosodic skills in children and adults. Results revealed nine test tools that assess prosody skills, namely: PROP (Crystal, 1982), PVSP (Prosody Voice Screening Profile – Shriberg, Kwiatkowski, Rasmussen, Lof & Miller, 1992),PEPS–C, Perception of Prosody Assessment Tool(PPAT; Klieve, 1998), Minnesota Tests of Affective Processing (MNTAP; Lai, Hughes, & Shapiro, 1991), Diagnostic Analysis of Non Verbal Accuracy (DANVA; Nowicki & Duke,1994), Aprosodia Battery (Ross, Thompson, & Yenkoshy, 1997), Florida Affect Battery (FAB; Bowers, Blonder, & Heilman, 1999), and Advanced Clinical Solutions (ACS; Pearson, 2009).

The most widely used battery of prosody perception available for purchase by the public is the Profiling Elements of Prosodic Systems—Children test, or PEPS-C (Peppé& McCann, 2003). This test assesses the perception and production of four prosody aspects namely, affect, phrase structure, focus, and interaction. The test has been successfully used to study the relationship between reading ability and prosody perception in typically developing children (Lochrin, Arciuli, & Sharma, 2015), and prosody perception in children with specific language impairment, dyslexia, and ASD (Wells & Pepé, 2003; Jarvinen-Pasley, Peppé, King-Smith, & Heaton, 2008; Marshall, Harcourt-Brown, Ramus, & van der Lely, 2009).

In the Indian context, preliminary attempts were made to understand aspects of prosody skills. Initial study in Indian context to investigate prosody skills is by Manjula (1997), wherein aspects of Kannada stress and intonation in WH and Y-N interrogatives were studied. Results revealed peak fundamental frequency and mean duration were greater for stressed syllables and rising intonation patterns for Y-N interrogatives and falling patterns for the WH questions. Priyanka (2018) studied speech rhythm patterns of spoken English in native Kannada-speaking adults who were proficient in English. Results revealed mora-timed rhythm in males and speech rhythm was unclassified in females. Studies report prosody deficits in clinical population, namely in persons with aphasia (Rohini, 2006), hearing impairment (Savithri, Johnsirani & Ruchi, 2008) and persons with stuttering (Santosh & Sahana, 2012).

Few attempts have been made towards the assessment of suprasegmental skills in Kannada. Mohan Natarajan (1991) developed a Synthetic test for intonation patterns, wherein 30 intonation patterns (such as gradual rise and fall, steep rise - gradual fall, gradual rise - steep fall to name a few) were synthesized with the target syllable /ba/. The study was carried out among 4-8 year old Kannada speaking typically developing children and results revealed a developmental trend for intonation perception and production. Synthetic test of rhythm was developed by Jayanthi ray (1993) and was administered among 40 Kannada speaking typically developing children in the age range of 2.6-6.6 years. The test included 17 tokens of rhythm patterns with frequency and intensity synthetic modifications and results revealed that with increasing age, the children's performance was better.

Divya (2005) had developed a manual in Kannada language for treatment of intonation in children with speech and language disorders. The manual has subsections for assessing and treating pitch height, pitch variation, pitch contour, nucleus in a pitch contrast and emotive intonation. All the subsections consist different modules for detection, discrimination and identification of intonation patterns. Akhila (2016) investigated the perception of affective prosody in Malayalam-speaking children with Specific Language Impairment (CwSLI) between 5-7 years of age. The results indicated that children with SLI made relatively more errors in identification of emotions than typically developing children. It was found that CwSLI identified the emotions *happy* and *sad* relatively better than other emotions such as anger, surprise and disgust in all the stimuli sentences.

The current study attempts to develop a preliminary assessment tool for prosody skills among Kannada-speaking children. Kannada is a Dravidian language which is characterized by a syllable-timed speech rhythm according to synthesis based studies (Patil, Shah, Patel & Sailor, 2013; Hemakumar & Punitha, 2014; Sai Geetha & Muralidhara, 2017). In normal sentential utterances, the speaker develops a rhythm of stressed and unstressed syllables. Certain languages (e.g., English) are called stress-timed because stressed syllables tend to occur at regular time intervals. Kannada was majorly reported to have mora-timed rhythm (Savithri, Jayaram, Kedarnath & Sanjay, 2006). Mora-timed rhythm is more similar to syllable -timed rhythm than stress-timed rhythm (Savithri et al., 2007). Many studies involving speech synthesis and automatic speech recognition have syllable units as their target since syllables are the basic unit of speech production and perception (Patil, Shah, Patel & Sailor, 2013; Hemakumar & Punitha, 2014; Sai Geetha & Muralidhara, 2017). The Load theory (Lavie, Hirst, de Fockert & Viding, 2004; Lavie, 1995, 2000) suggests that increased cognitive load in the input information leads to the recruitment of additional higher-order neural networks. Drawing inputs from the Load theory, the ability to perceive and produce lexical and contrastive stress increases the cognitive load in processing the stimuli or a speaker’s utterance (as the task involves processing lexical, suprasegmental and the link between lexical-suprasegmental information). The current study intended to assess prosody skills at syllable level as it is less cognitively taxing and syllables are considered as the basic unit of speech perception and production.

1.5 Need for the study

Prosody aids to convey attitudes and emotions (affective prosody), identify statement-question contrasts, differentiate word boundaries (grammatical prosody), emphasize relevant and new information and pragmatic aspects (pragmatic prosody) of speech. It is important to know how children comprehend different prosodic functions during communication development and the degree of variability that might be expected within and across age groups. Prosody captures the rhythm of speech, and the stress patterns within that rhythm, which provide emphasis and redundancy that facilitate understanding of the spoken message. Lack of rhythm results in speech breakdown leads to 'dysprosodia'. The sense of rhythm is poorly developed or is impaired in the hearing impaired, stuttering, cluttering, dysarthria, aphasia and verbal apraxia. (Schlanger, Schlanger, & Gerstman, 1976; Parkhurst, 1978; Starkweather, 1987). These populations need to be evaluated and rehabilitated for which the clinicians need to know about the development of rhythm. Additionally, speakers use dynamic pitch contours and intonational phrase boundaries to provide useful parsing information to the listener. It is important that a speech-language pathologist know about the development of suprasegmental aspects of speech, as children with communication deficits demonstrate unintelligible speech that is difficult to evaluate solely by segmental tests/segmental development norms. The above is necessary because in these population errors of rhythm and intonation also affect the intelligibility of their speech (Hudgins, 1946; John & Howarth, 1965; Hood & Dixon, 1969; Reilly, 1979; Stromberg & Levitt, 1979).

Studies suggest impaired prosody skills in children with communication disorders such as specific language impairment, learning disability, hearing impairment, autism spectrum disorder, etc, where both language and prosody skills are impaired. In the Indian context there are no published reports on norms for development of all three domains of prosody skills, which can be used to identify children at risk for communication disorders such as specific language impairment and learning disability. There is a strong need to develop assessment tools for prosody which is linguistically and culturally appropriate for an Indian context. Such investigations especially in the Indian scenario are needed to develop training materials for improving the acceptability of speech of children with communication disorders.

1.6 Aim

To develop a test for assessment of prosody skills in children and to establish norms for facilitating early identification of children with SLI

1.7 Objectives

1) To develop a test for assessing prosodic skills (stress, intonation & rhythm) in children.
2) To establish norms for prosodic skills for native Kannada speaking children using the developed test in the age range of 5-8 years.
3) To validate the developed test on age-matched children with SLI (CwSLI).

CHAPTER II

REVIEW OF LITERATURE

Prosody or the "melody of speech" is a complex and relatively unexplored aspect of linguistic function. Monrad-Krohn (1963) called it the "third element of speech," in addition to syntax and semantics. Prosody has variety of functions such as signaling the speaker's emotional state, marking syntactic structure and cues for turn-taking during conversations. Rhythm, timbre, contour, speech melody, stress and pauses are the components of prosody. Assessment of prosody is vital as majority of the communication disorders are characterized by impairment in prosody aspects. The literature reviewed is discussed under the following headings:

2.1. International status on assessment of prosody

2.2. National status on prosody assessment

2.2.1. Prosody skills in typical population

2.2.2. Prosody skills in clinical population (with communication disorders)

2.2.3. Assessment and training of prosody skills

2.3.Studies linking prosody and language skills

2.1. International status on assessment of prosody

Majority of the published reports on assessment of prosody are on English. The following section consists of details about prosody tests which have been primarily developed in English. Studies investigating prosody skills in other languages have adapted (only in case of meaningful stimuli; at syllable and sound level prosody is language-free) the developed tests in English for it to be linguistically and culturally appropriate for their specific population.

The tests developed for assessing prosody skills are as follows:

1. Tennesse Test of Rhythm and Intonation Patterns (T-TRIP; Koike and Asp, 1981)
2. Prosody Profile (PROP; Crystal, 1982)
3. Prosody-Voice Screening Profile (PVSP) (Shriberg, Kwiatkowski & Rasmussen, 1990)
4. Minnesota Test of Affective Processing (MNTAP; Lai, Hughes and Shapiro, 1991)
5. Diagnostic Analysis of Non-Verbal Accuracy (DANVA; Nowicki & Duke, 1994)
6. Aprosodia battery (Ross, Thompson & Yenkoshy, 1997)
7. Perception of Prosody Assessment Tool (PPAT; Klieve, 1998)
8. Florida Affect Battery (FAB; Bowers, Blonder and Heilman, 1999)
9. Profiling Elements of Prosodic Systems—Child version(PEPS-C; Peppe´& McCann, 2003)
10. Advanced Clinical Solutions (ACS; published by Pearson, 2009)

The above developed tests are reviewed in detail in the following sections.

1. Koike and Asp (1981) developed the ***Tennessee Test of Rhythm and Intonation Patterns (T-TRIP)***. The tool consists of three sections with 25 items to assess rhythm and intonation. The test items consist of the nonsense syllable /ma/ in 25 different rhythm and intonation patterns. The syllable /ma/ was uttered and audio recorded with 25 different rhythm and intonation patterns. The speaker was a young adult male with extensive phonetic training, who had undergone several practice trials prior to the

recording. Out of the 25 items, the first 17 test items assess rhythm while items 18 to 25 evaluate intonation skills. Items 1 to 14 constitute Part I of the test which includes 2 to 6 (minimum to maximum) /ma/ syllables that varied in stress (stressed/unstressed) and tempo (slow/fast). In Part II, the tempo of the stimuli increased in the range of 1 to 3 syllables per beat, accounting for 3 to 9 syllables (minimum to maximum) in test items 15, 16 and 17. Part III consisted of test items 18 to 25 which had various intonation patterns. The different intonation patterns used in the section were, namely, fast fall, fast rise, fast rise-fast fall, fast-fall-fast rise, gradual fall, gradual rise, gradual rise-gradual fall and gradual fall-gradual rise patterns.

The participants of the study consisted of 20 typically developing children with 10 participants from 3.0-3.11 age range and 10 participants from 5.0-5.11 years of age. All the subjects were exposed to the training session prior to the actual testing. 5 stimuli tokens from the 25-items were used for training purpose and the participants were trained in order to familiarize them with the task and method of response. In the experimental testing each stimuli token were presented twice to the participants and they were instructed to verbally imitate the perceived stimuli. On completion of the experimental testing, the responses of the participants were given to three judges who were instructed to listen and judge each response as correct or incorrect, in comparison to the stimulus. A participant’s response was considered as correct if at least two listeners judged it as correct. The criteria for a correct response included the appropriate use of stress, intonation, tempo and number of syllables in the stimuli. The results suggested that five-year old children scored significantly higher than the three-year old children. T-TRIP was concluded to be sensitive to differentiate the performance in children of different age groups.

2. ***Prosody Profile (PROP;*** Crystal, 1982) was developed in order to determine the expressive prosodic patterns present in an individual's speech. The aspects of prosody tested in the tool include pitch, tempo and stress. The tool consists of 4 subtests namely, intonation, tempo (phrasing), stress (phrasal) and strategies for producing stress. A connected speech sample is obtained from the participants and the subtests are administered. The responses are documented in detail with respect to pitch, tempo and stress. The test was made for clinical use and the guidelines cum exemplars of impaired prosody are provided by the author. The tool is not standardized and was developed for use among children and adults of all ages.

3. The ***Prosody-Voice Screening Profile (PVSP)*** (Shriberg, Kwiatkowski & Rasmussen, 1990) is a perceptual tool to evaluate an individual's prosody and voice in conversational speech. The tool quantifies a speaker based on seven suprasegmental domains such as phrasing, rate, stress, loudness, pitch, laryngeal quality and resonance. The test can be used for both children and adults in the age range of 3-81 years. Once a connected speech sample is obtained, the investigator codes each domain and a profile of prosody and voice is gathered. There are a total of 16 prosody codes assigned for each domain, such as 1 - Appropriate prosody, 2 - sound/syllable repetition, 12- fast/ acceleration, 14 - reduced/ equal stress and so on. The domains of voice such as pitch, loudness and quality are profiled using 20 various voice codes (such as 1-Appropriate, 23- Breathy, etc.,). The administration of PVSP requires prior training for the examiner in order to determine the prosody and voice characteristics of a speaker, based on a set of auditory-perceptual criteria through the use of codes. The tool majorly evaluates expressive prosody and is intended for clinical use.

4. Lai, Hughes and Shapiro (1991) developed the ***Minnesota Test of Affective Processing (MNTAP)***. Affective processing forms a vital part of prosody as emotions are conveyed by different intonation patterns. The tool evaluates the capability to recognize and differentiate between various affective stimuli presented through auditory and visual modalities. There are 16 subtests (4 auditory & 12 visual tests) namely, training tasks, inverted faces, 2 tasks of identity matching, faces teaching, affect match, affect naming, affect choice, gesture recognition, localization memory, face and object recognition memory, sequential face pairs memory and 4 auditory receptive subtests. The developed test was administered on 67 children with Attention Deficit Hyperactivity Disorder (ADHD) and 38 controls in the age range of 6 to 11 years. The results revealed developmental trend in all the tasks, especially in tasks such as inverted faces, face and object recognition memory and localization memory tasks. Children with ADHD performed poorer in all the tasks compared to their controls due to impaired affective processing and inattention. The MNTAP is mainly a receptive test which evaluates affective processing and takes a minimum of 2 hours to administer. The administration process is time consuming and the availability of the computerized tasks is restricted.

5. The ***Diagnostic Analysis of Non-Verbal Accuracy (DANVA***; Nowicki & Duke, 1994) was developed to assess the perception of facial expression, paralanguage (emotional prosody) and comprehension of body postures in both children and adults. The computerized tool consists of 5 subtests (2 face tests, 2 paralanguage tests & 1 posture test) namely, adult facial (AF) expressions, child facial (CF) expressions, adult paralanguage (AP), child paralanguage (CP) and adult postures. It is a picture selection test, wherein the participants should identify the appropriate pictures with four alternative forced choice response paradigm. The tool can be administered on

both children and adults in the age range of 3 to 99 years. The test administration duration is about 60 minutes and the normative data in terms of mean and standard deviation are available for all age ranges indicating a developmental trajectory for emotional prosody.

6. The ***Aprosodia battery*** (Ross, Thompson & Yenkoshy, 1997) was developed to examine the reception and expression of affective prosody in adults with acquired neurological disorders. The test battery consists of 4 subtests (2 receptive and 2 expressive) namely, affective - prosodic repetition, spontaneous affective - prosodic production, affective - prosodic comprehension and affective - prosodic discrimination. The various affective categories which are tested include happy, sad, angry, surprise, neutral and disinterested. The developed tool was administered on 22 brain-damaged subjects and 16 control subjects with the mean age of 49.4 years. The results revealed that persons with brain-damage performed poorly in all prosody tasks compared to controls. The test administration is about 60 minutes and can be used for adults above the age of 17 years.

7. The ***Perception of Prosody Assessment Tool (PPAT***; Klieve, 1998) was developed to evaluate prosodic perception in children with hearing impairment using cochlear implants in the age range of 7-12 years. There are 6 subtests namely, within linguistic context, apart from linguistic context, tone and affect, grammatical class, stress and compound words. The aspects of prosody tested include pitch, duration, intensity, grammatical aspects, affective (happy, sad, angry and sarcastic) and pragmatics. The test was administered on 6 children with hearing impairment using cochlear implants in the age range of 7-12 years. The duration of the test administration for each participant was about 60 minutes. The results revealed that children with hearing

impairment using cochlear implants perceived prosodic cues of duration, intensity and pitch better in the *'apart from linguistic context'* subtest than in the *'within linguistic context'* subtest. The performances of the participants were poor in other tasks. The developed tool is majorly a receptive prosody test for a clinical population (*hearing impairment*). It was mainly developed for research use and administration of the tool takes about 60 minutes.

8. Bowers, Blonder and Heilman (1999) developed the ***Florida Affect Battery (FAB)*** to evaluate the perception of affective prosody in adults with neurologic or psychiatric disorders. The tool consists of 11 subtests (5 facial, 4 prosodic & 2 cross-modal (facial-prosodic)) such as the facial affect discrimination, facial identity discrimination, facial affect selection, facial affect naming, facial affect matching, emotional discrimination, non-emotional discrimination, emotion naming, conflicting emotional prosody, matching emotional prosody to an emotional face and matching the emotional face to an emotional prosody. The developed test was administered on 164 participants in the age range of 17-85 years and normative values are provided for young adults (17-30 years), middle age adults (71-84 years), older adults (61-70 years) and elderly adults (71-84 years) for each of the 11 subtests. The test administration duration is about 60 minutes and the tool is developed to assess receptive prosody skills.

9. ***PEPS-C (Profiling Elements of Prosodic Systems—Child version***; Peppe´& McCann, 2003)was developed in response to a need for comprehensive prosody assessment. PEPS-C was modified from PEPS (Profiling Elements in Prosodic Systems, Peppe´, 1998), a test for adults which was developed and administered on 90 Southern British English (SBE) people aged 18–52 years. The tool was subsequently

administered on persons with aphasia (Peppe´, 1998). PEPS-C was devised and administered on 120 SBE children in the age range of 5–14 years. Results revealed higher scores with increase in age among the participants, thereby signifying the development of prosody with age. PEPS-C evaluates the ability to comprehend and express prosody with respect to *form level processing* (auditory discrimination and voice skills required) and *function level processing* (extent of communication affected due to impaired prosody). This approach resulted in four types of tasks: input form, output form, input function and output function. The test mainly taps into three domains namely, pragmatic, emotional and linguistic. Four core areas are assessed which include interaction, affect, chunking and focus.

In the interaction section various intonation patterns are assessed in a conversation sample. For example, a sharp rise at sentence boundaries indicate a request for repetition, while a falling pattern at sentence boundaries indicates understanding of what the speaker has said. The input and output functions are assessed in terms of strength and weakness. For instance, an input weakness would indicate little or no understanding of the concept that different intonation patterns have varying meanings. Output weakness suggests that the participant uses a single intonation pattern throughout the conversation or uses various intonation patterns with inconsistency.

In a similar way the other three core areas such as affect, chunking and focus are assessed. Affect refers to the way of conveying attitude, emotion or mood by intonation. Examples of tasks involved include expressing liking for a food item or drink as indicated by different intonation on a single syllable or word. Chunking refers to the grouping of words with respect to prosody, thereby generating syntactic divisions. One example used in the test to demonstrate chunking is the distinction between simple and compound nouns in a list (in the phrase ‘cream buns and jam’, the words ‘cream’ and ‘buns’ can be two simple nouns or one compound noun, depending on the presence or absence of a prosodic boundary after the word ‘cream’).Focus refers to the use of prosodic prominence in the significant part of an utterance. PEPS-C covers understanding and expression of variation in accent-placement at syllable and word levels.

The PEPS-C tool consists of 16 items in each of the four core areas tested. Each item has multiple (three- or four-way) choices to facilitate the alternative interpretations to be applied to one stimulus. It is a computerized test wherein the participants are instructed to point to the appropriate picture displayed on screen or to give a verbal response. Overall the tool provides comprehensive information about prosody skills by tapping into three domains namely, pragmatic, emotional and linguistic.

10. The ***Advanced Clinical Solutions (ACS)*** was published by Pearson in the year 2009. The tool was developed to assess social functioning deficits among adults in the age range of 16-70 years. The tool consists of 4 subtests, namely, facial affect recognition, affect naming, prosody-face matching and prosody-pair matching. The normative data is available from 800 participants aged 16-70 years. It is a receptive prosody tool which takes about 30-45 minutes to administer on participants and it is intended for both clinical and research usage.

Kalathottukaren, Purdy and Ballard (2015) reviewed the assessment tools that have been developed to assess prosodic skills in children and adults and to evaluate the clinical utility of the tools. The authors identified the available tools by using four online databases: Science Direct, SCOPUS, Web of Science, and PsycINFO. They also conducted manual searches of the bibliographies of published reviews and articles and contacted test developers to gain basic information regarding the tools. Results revealed nine test tools that assess prosody skills, namely: Prosody Profile (PROP; Crystal, 1982), PVSP (Prosody Voice Screening Profile – Shriberg, Kwiatkowski, Rasmussen, Lof & Miller, 1992), Profiling Elements of Prosodic Systems—Child version (PEPS–C; Peppe´& McCann, 2003), Perception of Prosody Assessment Tool(PPAT; Klieve, 1998), Minnesota Tests of Affective Processing (MNTAP; Lai, Hughes, & Shapiro, 1991), Diagnostic Analysis of Non Verbal Accuracy (DANVA; Nowicki & Duke,1984), Aprosodia Battery (Ross, Thompson, & Yenkoshy, 1997), Florida Affect Battery (FAB; Bowers, Blonder, & Heilman, 1999), and Advanced Clinical Solutions (ACS; Pearson, 2009).

From the identified test tools, norms are available for PEPS-C, DANVA 2, ACS, FAB and PVSP. Normative data for the PPAT, Aprosodia Battery, and MNTAP were obtained from comparative studies that evaluate the clinical and control groups. Only four tools (PROP, PVSP, PEPS–C and PPAT) are available that focus on two or more aspects of prosody. Among these, the PROP and PVSP require a high level of expertise to transcribe prosodic elements and are time consuming. The PPAT was originally developed for use with children using cochlear implants and therefore focuses only on assessing receptive prosody skills. The PPAT may work well with children with hearing loss, but limited empirical data are available. The PEPS–C was concluded to be a good choice in clinical practice as it covers a wide age range and is easy to administer and score. The authors conclude with the need to continue to develop and test tools for the effective and comprehensive assessment of prosodic skills in children and adults.

Kalathottukaren and Purdy (2017) studied prosody perception in typically developing school-age children. The participants were native speakers of New Zealand-English. Four receptive prosody subtests of the Profiling Elements of Prosody in Speech-Communication (PEPS-C) and the Child Paralanguage subtest of Diagnostic Analysis of Non Verbal Accuracy (DANVA) were administered to 45 children who were divided into three age groups, with mean ages 7.84, 10.13, and 11.90 years. The four receptive subtests of PEPS-C were namely, turn-end (interaction), affect, chunking and contrastive stress (focus) reception. The pre-recorded stimuli of PEPS-C were presented using a laptop through loudspeaker at a comfortable level in the range of 65-75dB SPL. The computer display screen of PEPS-C contains cartoon pictures and the children were instructed to point to the appropriate items or to give a verbal response.

The PEPS-C results showed that participants below 7 years of age performed significantly poorer than participants above 7 years on chunking and contrastive stress reception tasks, indicating a developmental trend. The reduced standard deviation scores and narrow ranges of scores obtained by9-12 year olds compared to the youngest group are also indicative of the age-related improvements. Overall the results indicate that much of the age-related changes in prosody perception occur between 7 and 9 years, especially in the 7-8 years age range. The DANVA test measures the competence in affect recognition by using voice tone (affective prosody) and reading facial expressions. The study used *child paralanguage* subtest to evaluate emotion recognition with voice. The results showed that 7-8 year old children made more errors, followed by 9-10 year olds, and least number of errors was made by 11-12 year olds. These results suggest a developmental trend in affective prosody perception abilities in children, however there was no statistical significance.

2.2. National status on prosody assessment

2.2.1. Prosody skills in typical population

In the Indian scenario studies were initially carried out to understand prosody skills and the major focus was on speech rhythm. An early study in the Indian context to investigate prosody skills is by Manjula (1979) to determine the intonation patterns in Kannada sentences and to analyze the responses of listeners to the emotional expressions of speakers. Thirty three typical native Kannada speaking adults (12 females and 21 males) in the age range of 19-35 years participated in the study. A total of 32 Kannada sentences were developed indicating both primary (joy, anger and grief) and secondary (neutral, jealousy, surprise, frustration, fear, worry) emotions. All stimuli sentences were uttered by two native Kannada speaking typical adults (1 male and 1 female) and were audio recorded. The recorded samples were perceptually analysed and it was found that there is a final fall in the intonation patterns of sentences with emotions such as jealousy, frustration and worry. The sentences with emotions such as surprise and joy consisted of rising intonation patterns. There was a final rise or fall in the sentences having the emotions fear and anger. In sentences with grief, the heights of rise and fall were perceived to be less and more gradual in nature. The neutral sentences also had rise and falls, however the degree of rise and falls were perceived to be less. Following the perceptual analysis, the recorded sentences were presented to all 33 participants and they were instructed to identify the perceived emotions. Results revealed that the listeners had higher identification percentage for the emotions anger, jealousy, surprise and grief when the stimuli sentences were spoken by male speaker. For other emotions such as fear, frustration, worry, joy and neutral, listeners scored higher when the stimuli sentences were spoken by a female speaker.

Manjula (1997) studied the perceptual and acoustic correlates of intonation and stress in WH (questions -what, where, etc) and Y-N (Yes/No) interrogatives in Kannada language. The participants of the study were 60 typical adults in the age range of 20-45 years, who were native speakers of standard Kannada dialect. Participants were chosen and randomly assigned in to two groups as ‘speakers’ and ‘listeners’. A picture from the ‘Visual closure’ sub-test of *Illinois test of Psycholinguistic abilities* (Kirk & McCarthy, 1968) was chosen as the stimuli. The speakers were instructed to ask ‘Wh’ and ‘Yes-No’ questions about a picture stimuli as naturally as possible to the listener. The conversation/ discourse by the speakers/ listeners were recorded and only the questions asked by the speakers were selected and analyzed. The intonation contours of the interrogatives uttered by the participants were primarily rising patterns (low to high frequency). With respect to stress, it was observed that peak fundamental frequency and mean duration were greater for stressed syllables. It was also observed that percentage of stress was more in geminates, followed by consonant clusters and morphemes in ‘Wh’ and ‘Yes-No’ interrogatives.

Savithri, Jayaram, Kedarnath and Sanjay (2006) investigated rhythm in two etymologically unrelated languages namely, Indo- Aryan (Hindi) and Dravidian (Kannada). Twenty typical adult speakers (10 males and 10 females) of each language in the age range of 18-25 years participated in the study. Spontaneous speech samples were collected and the participants were also instructed to read a 1000-word passage. The recorded samples were analyzed using nPVI and rPVI measures. The results revealed Kannada to be a mora-timed language and Hindi as a syllable-timed language.

Savithri, Maharani, Sanjay and Deepa (2007) investigated speech rhythm in 12 Indian languages, namely, Assamese, Bengali, Gujarati, Kashmiri, Marathi, Oriya, Punjabi, Rajasthani, Kodava, Malayalam, Tamil and Telugu. Twenty typical native speakers with 10 males and 10 females in all languages in the age range of 18-35 years participated in the study. Participants were instructed to read a 1000-word passage (in their respective languages except for Kodava and Rajasthani as written script was unavailable) and describe a picture (developed to elicit spontaneous speech). The recorded samples were analysed using nPVI and rPVI measures. Results revealed mora-timed speech rhythm in Assamese, Punjabi and Telugu; Bengali, Malayalam, Kodava, Tamil and Kashmiri as syllable-timed languages. Rajasthani, Gujarathi and Kannada were unclassified since they did not fall in any of the category (syllable/ stress/ mora-timed speech rhythm) of the Pairwise Variability Index of Grabe's (2002) chart. Marathi and Oriya were found to have syllable-timed rhythm in reading task and mora-timed rhythm in speech task.

Savithri, Sreedevi and Kavya (2009a) investigated speech rhythm in typically developing native Kannada speaking children. Five children (males) in the age range of 8-9 years participated in the study. A five-minute speech sample of each child was elicited using pictures depicting simple stories and their responses were audio recorded. The vocalic and intervocalic durations were measured in the samples using Adobe Audition software and measures such as rPVI and nPVI were calculated for the samples. Results revealed rhythm to be unclassified in the participants. The authors attributed this finding to ongoing language acquisition in children and predominant use of vocalic patterns, which is different from adult usage.

Savithri, Sreedevi and Kavya (2009b) explored the differences in the type of speech rhythm between typically developing male and female native Kannada speaking children. Ten children (5 boys & 5 girls) in the age range of 8-9 years participated in the study. A five-minute speech sample of each child was elicited using pictures depicting simple stories and responses were audio recorded. The vocalic and intervocalic durations were measures in the samples using Adobe Audition software. Rhythm metrics such as nPVI and rPVI were computed and results revealed rhythm patterns to be unclassified with no effect of gender.

Savithri, Sreedevi, Jayakumar and Kavya (2010) investigated the development of speech rhythm in typically developing native Kannada speaking children. Fifteen boys divided into three age groups (4-5, 8-9 & 11-12 years) equally distributed in number were considered as participants. A five-minute narration sample was elicited using cartoons or Panchatantra pictures. The vocalic and intervocalic durations were measured in samples using PRAAT software. Measures such as nPVI and rPVI were calculated and results revealed rhythmic patterns to be unclassified. It was also found that the participants begin to produce adult-like rhythm around 11-12 years of age.

Priyanka (2018) carried out a study to determine the speech rhythm pattern in native Kannada speakers speaking in English. The investigation included 20 literate adult native Kannada speakers (10 males and 10 females) in the age range of 20-30 years. All were proficient in English which was ensured using English proficiency scale (Ramya & Goswami, 2009). A reading task including 5 standardized sentences in English (White & Mattys, 2007) was used as the stimuli. The participants were instructed to read the sentences in their habitual manner and the utterances were audio recorded. Speech samples were transferred and Praat (5.1.14) software was used to measure vocalic and intervocalic durations. Rhythm metrics such as the nPVI and rPVI scores were calculated. Results revealed the rhythm class in male population as mora-timed and in females, it was unclassified.

2.2.2. Prosody skills in clinical population (with Communication Disorders)

In the Indian context, studies have investigated speech rhythm in individuals with communication disorders such as aphasia, hearing impairment and stuttering. Rohini (2006) investigated perception of stress in 100 participants with Cerebro Vascular Accident (CVA) (50 participants with right hemisphere damage and 50 participants with left hemisphere damage) and 50 typical controls. The participants were native speakers of Kannada and were in the age range of 21 to 80 years. The stimuli consisted of 25 meaningful two-word Kannada phrases with the first word being an adjective and latter being a noun. The stimuli phrase were uttered in a neutral tone by a female native Kannada speaker aged 42 years. Synthetic tokens of the phrases were generated using the ‘PATPLAY’ program of SSL Pro 2V2 software. The first word of the phrases was stressed (through synthetic modification) and the second word of the phrase remained unstressed (original). Binary tokens were created, wherein the original recorded phrase (Unstressed-Unstressed) was paired with the synthetic phrase (Stressed- Unstressed). The participants were instructed to listen to each phrase and indicate whether two phrases in the pair were ‘same’ or ‘different’ by marking on a binary forced choice response sheet under the respective category.

Results revealed that the percent discrimination scores were higher for phrase pairs altered with combination of acoustic parameters (frequency, intensity and duration together) compared to those altered in single parameters (frequency/ intensity/ duration). Typical participants discriminated maximum number of phrase pairs when duration was altered compared to other single cue conditions such as intensity only or frequency only. Among participants with CVA, participants with left hemisphere damage had higher scores on altered frequency-duration and frequency-intensity-duration conditions compared to participants with right hemisphere damage.

Savithri, Johnsirani and Ruchi (2008) investigated speech rhythm in Kannada speaking children with hearing impairment. Twenty typically developing children (7 males & 13 females) and 20 children with hearing impairment (9 males & 11 females) in the age range of 5-10 years participated in the study. Picture description and story narration tasks were carried out and the responses of the participants were audio recorded. The recorded samples were analysed using rhythm measures such as nPVI and rPVI. Results revealed that the speech rhythm in typically developing children and children with hearing impairment remained unclassified. No gender differences were obtained for both groups.

Santosh and Sahana (2012) studied the rhythmic characteristics of speech of persons with stuttering (PWS). The study included 30 participants in the age range of 17 to 30 years with 15 PWS and 15 gender-matched typical adults. The participants were native speakers of Kannada language. Five short sentences were formulated in Kannada and were given to five native Kannada speakers for familiarity rating. The participants of the study were instructed to read the sentences in their habitual loudness and comfortable rate and the utterances were audio recorded. Rhythm metrics such as nPVI, rPVI, PerV, DeltaC, DeltaV, VarcoV and VarcoC were calculated. The mean values were higher for PWS than control speakers for the rhythm metrics signifying a high variability between successive vocalic and inter-vocalic segments of speech. The results indicate that AWS are less regular in timing of speech events compared to individuals without stuttering.

2.2.3. Assessment and training of prosody skills

When considering the area of prosody assessment, initial studies on Indian context available are by Mohan Natarajan (1991), who developed a synthetic test of intonation patterns. The study focused at understanding the performance of children aged 4-8 years, in imitating the synthesized intonation patterns. The study included 30 intonation patterns containing syllable /ba/, which were synthesized using the software developed by Voice and speech systems, Bangalore. The 30 intonation patterns included gradual rise, steep rise, gradual fall, steep fall, gradual rise and fall, steep rise and gradual fall, gradual rise and steep fall, steep rise and steep fall, gradual fall and along with gradual rise, gradual fall and steep rise, steep fall and gradual rise, steep fall and steep rise and so on. A total of four levels of frequency were identified with 230Hz being the base (Level 1), 160Hz being the level below the base (Level

2) and two levels - 300 Hz and 370Hz (Levels 3 & 4) were identified as above-base levels. The duration of each stimulus varied between 500 milliseconds to 1000 milliseconds depending on the target intonation pattern.

A total of 30 stimuli a token were generated and converted to analog signals and were audio recorded in a noise-free environment. A pilot study was carried out on 3 Speech Language Pathologists (SLPs), where they were instructed to listen to each stimulus and transcribe the stimuli tokens in terms of gradual rise / fall, steep rise/ fall and steady components. It was found that no single pattern (except the steady components) was identified in its level/ steepness by the three judges. Following the pilot study, the experiment was carried out on 20 Kannada-speaking typically developing children in the age range of 4-8 years. The participants were instructed to imitate the intonation patterns perceived during the testing phase. The synthesized stimuli and the imitated responses of the children were equated by 2 SLPs. The SLPs were instructed to mark each token as "S" (same) if a similarity existed between the stimuli and the imitated response, failing which it was marked as "D" (different). The results of the study revealed that the ability of children to imitate the intonation patterns increased from 4-7 years and declined thereafter. The decline in performance after 7 years was not discussed in the study. There were no gender differences in terms of imitation of the intonation patterns. This tool cannot be considered as a comprehensive test for children, since the adult listeners were not able to identify all the 30 intonation patterns.

Jayanthi Ray (1993) developed a synthetic test of rhythm. The test consists of 17 rhythm patterns, which were synthesized based on three parameters – change of intensity, change of fundamental frequency and change of both fundamental frequency and intensity. The target stimuli /ba/ of 500 ms duration and a sampling frequency of 8000 Hz, was synthesized using SSL, a software developed by Voice and speech systems, Bangalore. Rhythmic patterns starting from one foot to six feet were generated, following the method of re-iteration where stressed and unstressed syllables were concatenated. The original /ba/ syllable had a fundamental frequency of 140 Hz. Intensity was manipulated in 10 dB steps starting from 60 dB and F0 was manipulated in 5 Hz steps. The study involved testing the adults initially, for identification of rhythm patterns. A total of 17 stimuli were audio presented to 20 normal adults, who were instructed to imitate the same. The recorded imitations, were given a score of "1", if it matched with that of the stimuli and "0" was assigned for inappropriate imitations.

Results indicated that adults could imitate up to five feet and had difficulty in imitating six feet condition. The adult participants opined the five feet condition to be difficult for administration on children. Thus, while testing the children five and six feet stimuli tokens were deleted and the remaining stimuli were tested on 40 Kannada speaking typically developing children in the age range of 2.6-6.6 years. The children were instructed to imitate the rhythm patterns, which were recorded and scored, similar to the procedure done for adults. The results revealed that with increase in age, there was a linear increase in performance scores, depicting the developmental trend for rhythm in children. There was no gender differences noted in terms of the performance of the participants.

The manual for correction of Intonation in Kannada-speaking children with speech and language disorders (MCI-K) was developed by Divya (2005) for targeting correction of intonation. The manual can be used for Kannada speaking children above three years of expressive language age and with impaired intonation due to any speech and language or communication disorder. It can also be used for children with hearing impairment, provided they use an amplification device and should have undergone auditory training for a minimum of 6 months duration. The manual has 5 sections, namely pitch height, pitch variation, pitch contour, nucleus in a pitch contour and emotional sentences. Each section has 2 subsections, namely the perception of intonation and production of intonation. The perception of intonation section consists of 3 modules: A- Detection of intonation; B- Discrimination of intonation and C- Identification of intonation. The manual has both testing phase and training phase. In testing phase, each correct response should be scores as '1' and every incorrect response as '0'. A set criterion of 80% correct responses should be achieved, before proceeding to the next section or the subsection. The same scoring procedure and the 80% criterion is followed in the training phase. The developed manual was administered on three children with delayed speech and language for item validation. Subsequently, the instruction and the activities were modified in the manual. The developed manual can be validated on a larger group of children with communication disorders, which can shed light on sub grouping the disorders based on level of intonation deficits.

2.3. Studies linking prosody and language skills

Marshall, Brown, Ramus and Van der Lely (2009) studied the relationship between prosody and language skills in CwSLI and or dyslexia in the age range of 10-15 years. The study included 3 clinical groups, namely, SLI with dyslexia, SLI only and dyslexia only; and 3 control groups- 2 younger groups which were language and literacy matched and other group matched for chronological age. Profiling Elements of Prosodic Systems- Child version (PEPS-C) was administered to explore comprehension/ discrimination and production/ imitation of prosodic forms. Results revealed that most of CwSLI and or dyslexia performed well on tasks of auditory discrimination and imitation of prosodic forms. However, CwSLI had impairment in use of prosody to differentiate among linguistic structures, when compared to their age peers. The authors concluded that interaction between prosody and other language components such as syntax and pragmatics is affected in CwSLI and or dyslexia.

Also, prosody as an individual component may not be the core impairment in CwSLI and or dyslexia.

Corriveau and Goswami (2009) explored the expression of motor abilities among CwSLI on motor tasks, which require rhythmic processing and on motor tasks which lack a rhythmic component. The study included 63 children in the age range of 7-11 years, with 21 CwSLI (mean age- 10.2 years), 21 age-matched controls (mean age- 9.9 years) and 21 language ability-matched (mean age- 7.8 years). All children were subjected to tasks such as language test, word recall task, non-word repetition task and motor tasks. The motor tasks consisted of expressive rhythmic timing and Pegboard task. The expressive rhythmic timing (metronome) task was designed to compare rhythmic motor ability in paced settings (in time to metronome beeps) versus rhythmic motor ability in unpaced setting (continuing to tap with the previous pace, without metronome beeps). The experiment was carried out on laptop with sounds presented through headphones at 73 dBSPL. Each metronome speed was presented for 30 seconds (paced condition) followed by 30 seconds block of silence (unpaced condition).In both conditions, children were instructed to respond by pressing the "spacebar" key and the Inter-Tap-Intervals (ITIs) were calculated. For the Pegboard task, children were instructed to place pegs of 1 cm in diameter on pegboard holes with each hand individually and then with both hands together. Results of the study revealed that CwSLI were impaired in paced rhythmic tapping, but not equally impaired in tapping in unpaced condition (as other two groups also performed poor in unpaced condition). The authors opined that the severity of impairment in paced tapping was linked to language and literacy outcomes. In pegboard tasks, CwSLI were significantly poorer than other two groups. The rhythm deficits observed in CwSLI was attributed to impairment in rhythm perception and procedural memory.

Clement, Planchou, Béland, Motte, and Samson (2015) explored the singing abilities in CwSLI. The study included 8 CwSLI (2 females & 6 males) and 15 typically developing children (6 females & 9 males), who were matched for age, gender and non-verbal intelligence. The participants were in the age range of 8-12 years. The testing consisted of pitch-matching and melodic reproduction. The pitch-matching task involved the participants to sing back the audio musical notes (played with piano sound) on the syllable /la/. In the melodic reproduction task, the children listened to melody played with a piano sound. After each presentation, the children were instructed to sing it back on the /la/ syllable. A total of 6 melodies (2 familiar melodies associated with lyrics, 2 familiar melodies not associated with lyrics and 1 non-familiar melody) were presented as stimuli. Results revealed that CwSLI were significantly impaired in pitch-matching task and melodic reproduction task, when compared to age-matched peers. 30 adults rated the quality of sung productions of all the children on a continuous rating scale and it was found that CwSLI had obtained lower mean ratings than typically developing children.

Sallat and Jentschke (2015) studied the influence of music perception on language acquisition. The study included 29 children with SLI, with mean age of 64.2 months, 39 chronologically age-matched children with mean age of 63.6 months and 13 language-matched children with mean age of 51.3 months. The experiment consisted of 2-3 sessions for each individual, with the session duration f about 20-25 minutes. First session involved language screening of the participants using standardized German test tool for language screening for 3-5 years old children (SETK 3-5; Grimm, 2001); nonverbal intelligence was tested in second session. Remaining sessions consisted of evaluating music skills, which taped on key aspects, namely, melodic perception, melody recognition and rhythmic-melodic perception. The stimuli consisted of audio files (played on piano and recorded) containing the beginning phrases of nursery rhymes, which were presented through loudspeakers. The stimuli were presented on a laptop, using Presentation 0.76 software and the responses were recorded automatically. After each stimulus, the children were instructed to do button-press responses accordingly (one button-for correct musical phrase tone and a different button for incorrect musical phrase tone) in the melodic perception task. For the melody recognition task, one picture associated with the rhyme was displayed on the laptop; hence a total of 4 pictures were displayed on screen and 4 different audio file was presented. The children were instructed to respond through button press, by selecting the appropriate picture for the stimuli. The audio files were manipulated in 4 different ways (original tune of the rhymes; modified- same tune but played on a different note; tempo change-faster tempo) to check for accurate melody recognition. In all the above tasks, the accuracy of the task was calculated and all children responded through button-press method. The results suggested that CwSLI performed poorly in all tasks, suggesting a significant deficit in non-linguistic processing.

Mari, Scorpecci, Reali and Alatri (2016) studied the music identification skills of CwSLI and possible correlations between music identification skills and language abilities in SLI group. The study consisted of 30 CwSLI (21 males & 9 females) with mean age of 56+/-9 months and 23 age-matched children (11 males & 12 females) with mean age of 61+/-12 months. Melody identification and song identification scores were assessed. Songs used included 5 popular Italian tunes and the melodies were presented as piano versions. Both subtests were done by audiologists. In each task, 5 musical items were presented in random order through speakers. After the auditory stimuli, the children were instructed to point to the starring character among the 5 pictures presented on a laptop. For instance, the children are supposed to point to a picture with a *star*, for a rhyme like *twinkle-twinkle little star*. The children were not given any feedback regarding the correctness of responses and were instructed to guess correctness of responses, if they were not sure of the character. The results revealed that song and melody identification scores among SLI children were significantly lower than TD children. The authors also found significant correlation between song identification skills and chronological age in both groups.

Akhila (2016) investigated the perception of affective prosody in children with Specific Language Impairment (CwSLI) between 5-7 years of age. A total of 40 children (10 CwSLI and 30 typically developing children (TDC)) who were native speakers of the Malayalam language were recruited for the study. Three sets of stimuli were developed for the study. The first set (non-sense set) consisted of 25 sentences in Malayalam language that was syntactically correct but semantically anomalous. The second set (neutral set) consisted of 10 neutral sentences and the final set (emotional set) included 25 sentences imparting the target emotions (happy, sad, anger, surprise, disgust and neutral). The stimuli sentences were uttered and recorded by a female native speaker of Malayalam in a sound treated room in CSL software. The sentences were presented to each participant through headphones in a random order and a picture pointing task was employed. Six pictures of emoticons (symbols) depicting the targeted emotions were placed on card in random order. The participants were instructed to listen to the stimulus and then identify the emotion conveyed by pointing to the appropriate emotion on the card. The correct responses were scored as '1' and the incorrect as score '0'. The findings of the current study revealed that the performance of TDC for the emotions - happy, sad and anger was relatively better than that of surprise and disgust in all the sentence types. The performance of TDC on perception of emotional sentences was found to be better compared to perception of non-sense and neutral sentences. In addition, a positive age effect was also found in the TDC group with better performance of TDC in emotion identification for all the sentence sets. In the tasks on perception of emotional, non-sense and neutral sentences, TDC and CwSLI performed similarly on recognition of the emotions - happy and sad, whereas CwSLI were relatively poor on recognition of emotions such as anger, surprise and disgust.

To conclude, an everyday major challenge faced by Speech language pathologists (SLPs) is in assessment and intervention of prosodic skills of persons with communication disorders. There are very few test tools available to assess prosody compared to the availability of numerous tools for other speech and language components such as voice, resonance, semantics, syntax, articulation and phonology. Majority of the existing prosody tools are either narrow in their scope and usage or have not been validated. Studies have been carried out using experimental tasks to assess the development of prosody and impairments in prosody among typically developing children and in children with communication disorders such as CwSLI. A standardized and comprehensive prosody test battery which is easy to administer among children is warranted in the Indian context.

CHAPTER III

METHOD

The present study aimed to develop a tool for assessment of prosody skills in typically developing native Kannada speaking children and validate the same on children with Specific Language Impairment (CwSLI) aged between 5-8 years.

3.1. Study design

The study followed a cross-sectional research design under normative research.

3.2. Participants

The study included a total of 100 children consisting of 2 groups. Group I served as the control group, with 90 typically developing (45 boys and 45 girls) participants in the age range of 5-8 years (15 boys & 15 girls in each age group: 5-6; 6-7 & 7-8 years). The sample size was estimated based on the project time line of 1 year and did not involve the usage of any formula or a sample size estimation software based on an article. Random sampling was carried out to recruit the participants for the study. The participants were screened for any disability using WHO 10-disability screen questionnaire and were recruited from local schools. Group II included 10 age-matched children with SLI (CwSLI), who were selected based on Leonard's exclusionary criteria (1988) which is as follows:

1. Language abilities – Language test scores of at least -1.25 SD
2. Non-verbal IQ – 85 or higher
3. Hearing – Passes screening at conventional levels.
4. No recent episodes of otitis media with effusion.
5. No evidence of seizure disorder, cerebral palsy, brain lesions and not under medication for control of seizures.
6. No oral structural anomalies and developmentally appropriate oral-motor function.
7. No symptoms of impaired reciprocal social interaction or restriction of physical activities.

Group II participants were recruited from AIISH, with the clinical characteristics of

SLI (based on Leonard's exclusionary criteria) and were availing speech language therapy services.

3.3. Procedure

3.3.1. Phase I - Development of the Prosody Assessment Test

A synthetic test for prosody skills was developed using the Praat software (Version 5.4.09). This test consists of two major tasks, namely,

i. Verbal

ii. Non- Verbal

Each task consisted of three major sub tasks of prosody such as,

i. Stress

ii. Intonation &

iii. Rhythm

3.3.1.1. Stimuli preparation

The stimuli consisted of tokens which were synthesized using the manipulation and intensity tier options available in Praat - version 5.4.09 (Boersma & Weenink, 1992). The fundamental frequency, duration and intensity of the stimuli were altered one at a time and synthesized according to the task. There are two sections in the study: verbal and non-verbal to assess prosodic skills among children. In both verbal and non-verbal sections there are three tasks each to assess stress, intonation and rhythm.

i. Stress task

a) Verbal stress task

The syllable /ma/ audio recorded from an adult male speaker aged 20 years was used for synthetic modification of stress. A digital recorder (Olympus WS 100) was used to record the stimuli. Syllable /ma/ was chosen as bilabials are acquired early in the phonological development of children (Vihman & Croft, 2007; Sunny & Kumaraswamy, 2015). Nasal consonants are rich in static and dynamic cues. The nasal murmur has intense energy which can be perceived easily when compared to the acoustic cues of other consonants (Ohde, 1994). Hence, this makes the syllable /ma/ as an appropriate choice for the stimuli. The target syllable /ma/ was synthetically modified by increasing the value of one of the prosody markers, that is, either the fundamental frequency or intensity or duration. For the fundamental frequency and duration manipulations, the "Manipulate" option was used from Praat.

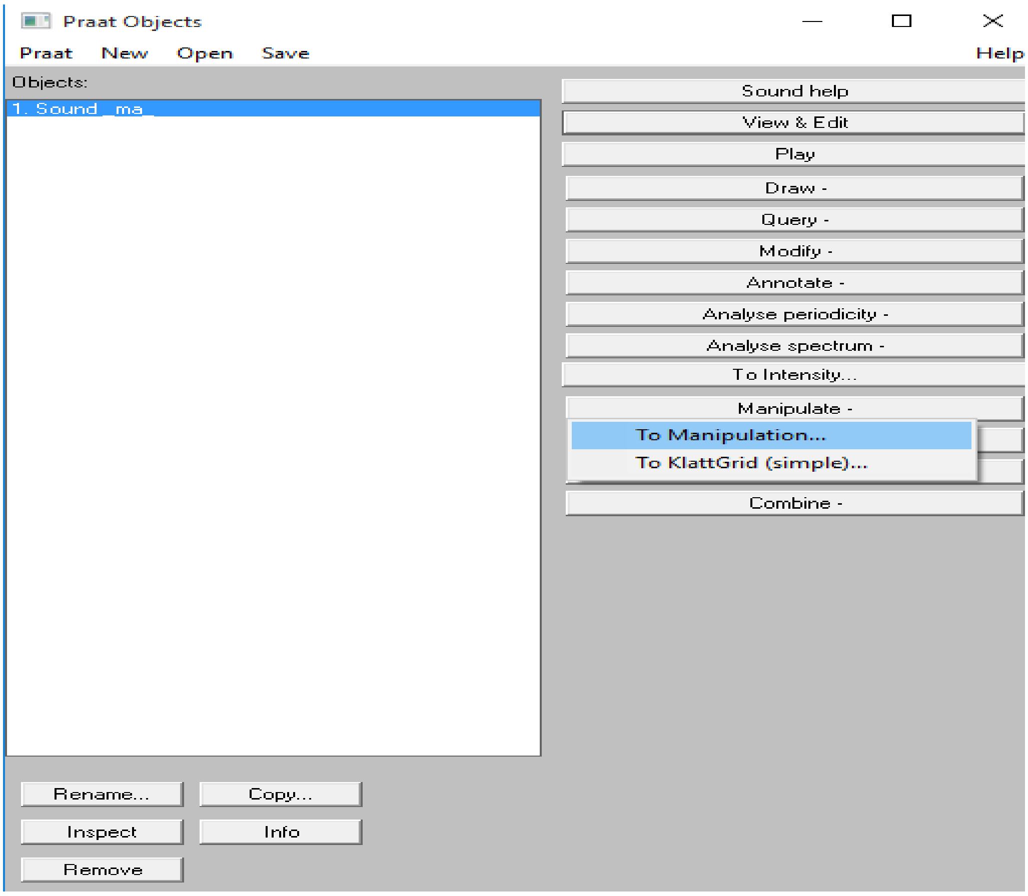


*Figure 3.1.*Screenshot of Manipulation window (both for frequency & duration parameter)

The following steps are followed (in Praat) for synthetic manipulation of stress for the frequency and duration parameters:

i. Open the recorded file in Praat menu.

ii. Select the file and go to "To manipulation" option. Select "Ok" for the default settings.

iii. Select the "Manipulation" option that appears on the Praat menu.

iv. Click "View & Edit" option by selecting the "Manipulation" file on the menu.

v. Select points (plot points) on the frequency manipulation window (refer to Figure 3.2) according to frequency/duration of interest. For example in Figure 3.2, points have been plotted for a frequency of 107Hz.

vi. The window will have options for both frequency and duration points, which can be plotted.

vii. After plotting the points, go to "File" and select "Publish re-synthesis" option.

viii. Save the new sound (target synthetic token) file that appeared on the menu.

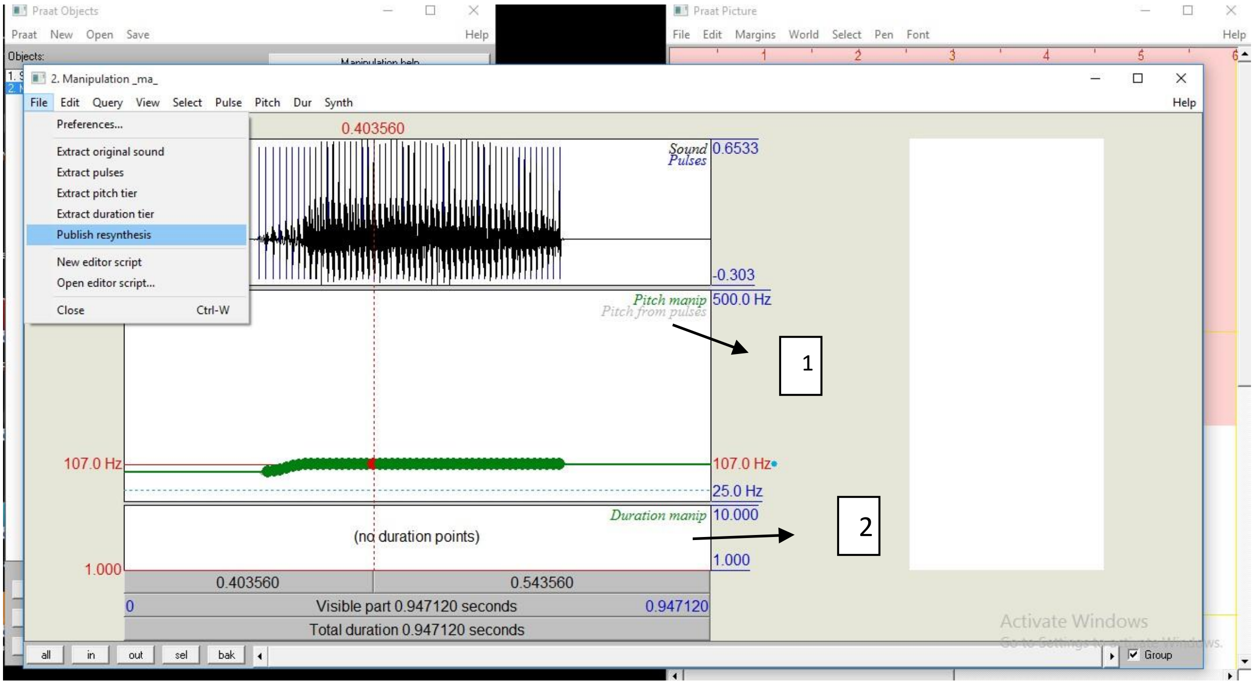


Figure 3.2. Screenshot of manipulation window {with options for both pitch (1) and duration manipulation (2)}

For the intensity manipulation, following steps are followed:

i. Select the recorded file and go to "To intensity" option (See Figure 3.3).

ii. Give "Ok" for the default settings.

iii. Select the intensity file that appears on the menu and go to "To intensity tier peaks" option.

iv. Select the "Intensity tier" file that appears on the menu and go to "View & Edit".

v. Select the intensity gain points in the manipulation window (See Figure 3.4).

vi. Select the recorded file on the menu and the gain points (intensity tier) file together and click "Multiply". Give "Ok" to the setting window, after removing the default selection. (Default: The sound will be scaled to 0.9 times).

vii. Save the final sound file that appears on the menu. Intensity can be modified further, according to the requirements, by changing (increasing/decreasing) the gain points on the manipulation window.

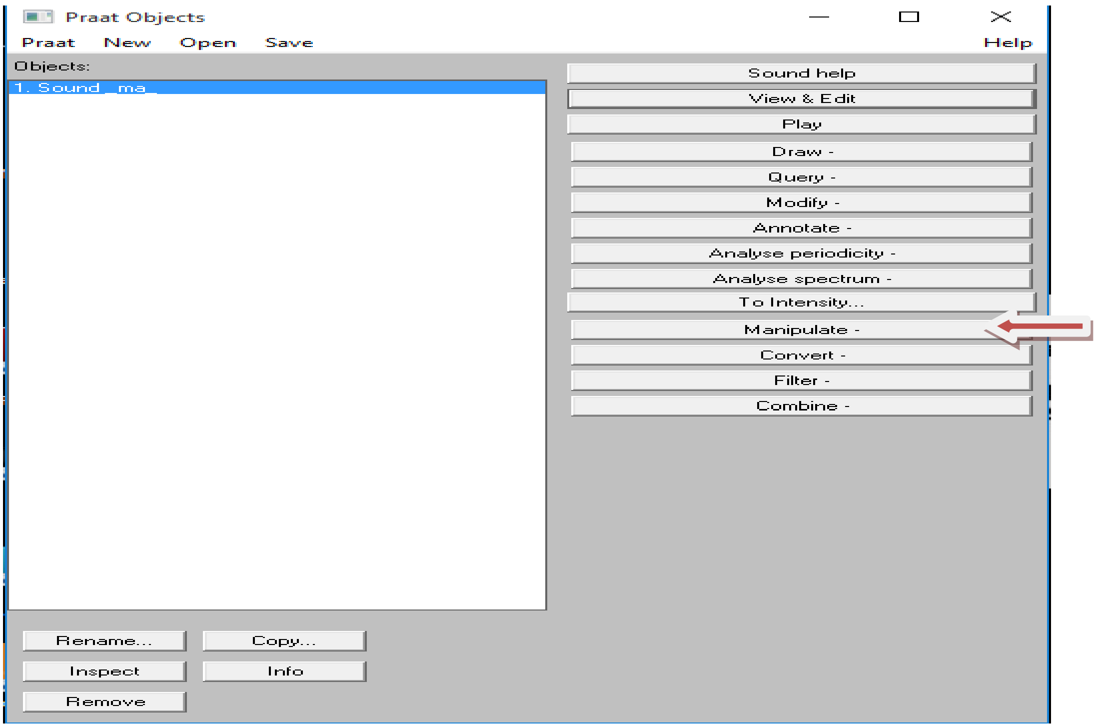


Figure 3.3. Screenshot of intensity manipulation option

Figure 3.4. Screenshot of plotting intensity gain points

As explained, three types of synthetic manipulations with frequency or duration or intensity were carried out to obtain the stressed tokens while the original recorded audio file (from 20-year old adult male speaker) was retained as the unstressed token. Each type (frequency or duration or intensity) of synthetic manipulation consisted of three binary combinations, namely,

- Stressed- Unstressed (S-US) (i.e, Synthetic token –Original token)
- Unstressed-Stressed (US-S) (i.e, Original token– Synthetic token)
- Unstressed-Unstressed (US-US) (i.e, Original token–Original token)

b) Non-verbal stress task

The task consisted of a series of tones played in an online virtual piano (https://virtualpiano.net/). Available since 2006, the Virtual piano is a free to use platform that enable us to play the piano through the computer keyboard. The stimulus (generated piano middle C note – with a frequency of 261.63 Hz) was recorded using a digital voice recorder and was converted as *.wav* file. The recorded piano tone was synthetically modified in terms of frequency or intensity or duration (as described in the verbal stress task) one at a time. A similar stimulus using piano notes was used by Sallat and Jentschke (2015) for melody recognition task. In the current study, each token consisted of various binary combinations similar to that of verbal stress task.

ii. Intonation task

a) Verbal intonation task

A phonation of vowel /a/ was recorded from a 20-year old male speaker. Pitch glides of vowel /a/ were synthesized in 4 different patterns {Falling (High-Low), Rising (Low-High), High steady (High-High) and Low steady (Low-Low)}. The "Manipulation" option as stated earlier was used from Praat to obtain the same. Vowel /a/ was chosen as the stimuli instead of syllable /ma/ since the intonation patterns of /a/ was perceived better than that of syllable /ma/ by adult listeners (Refer 3.2.2. for details). Also, it was easier to synthesize pitch glides in Praat and the glides were perceived much clearly when vowel /a/ was used than syllable /ma/.

b) Non-verbal intonation task

The stimuli consisted of tone glides, which were synthesized using Praat. In order to generate tone glides, pure tones were synthesized using the following steps:

i. Open Praat.
ii. Go to "New" in Praat menu.
iii. Go to "Sound.
iv. Select "Create sound as pure tone".
v. Give the duration and frequency of the tone accordingly.
vi. Select "Ok".

The generated pure tones were synthetically modified into tone glides, using the "Manipulation" option available in Praat. Similar to the verbal intonation task, pure tone pitch glides were synthesized in four different patterns (High steady, Low steady, Rising and Falling). The tokens were presented auditorily to the participants in a random order, accounting for a total of 8 tokens (four patterns presented twice randomly).

iii. Rhythm task

a) Verbal rhythm task

The verbal rhythm task was based on the Synthetic test of rhythm (developed by Jayanthi Ray, 1993) in the Indian context. The test was developed and used for 2.6-6.6 year old typically developing children. The test contains 17 tokens with syllable /ba/ (due to lack of accessibility to synthesize/manipulate the nasal sounds way back in 1993) with manipulations done in terms of fundamental frequency (F0) and intensity. The participants are supposed to imitate the tokens as perceived.

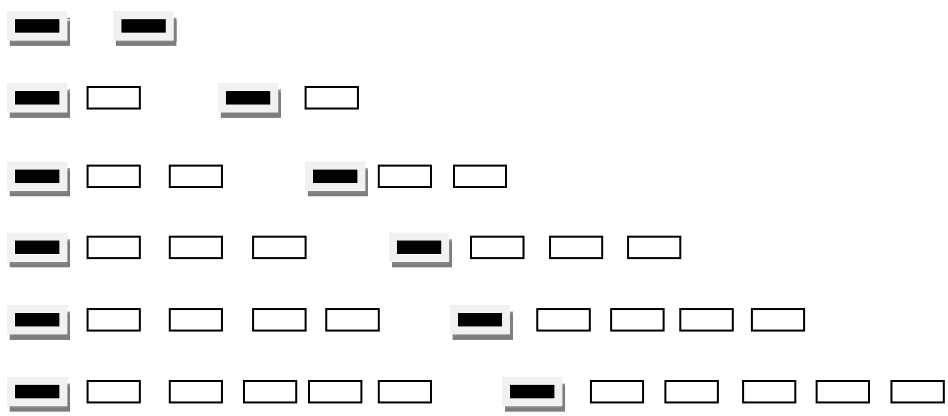
For the present study, similar manipulations were done using syllable /ma/. Syllable /ma/ was chosen as it is acquired earlier in phonological development. The current task was developed with the syllable /ma/ using manipulation option (fundamental frequency manipulation) and Intensity tier option (intensity manipulation). Only frequency and intensity parameters were considered for the verbal rhythm task. The developed stimuli material was used to assess rhythm in both Group I and Group II participants.

The method of reiteration was used to generate rhythmic patterns, which were based on a study by Savithri (1991). The earlier recorded syllable /ma/ from a 20-year old male speaker (*refer page 3*) was considered as the stimuli. Rhythmic patterns from one foot to six feet were synthetically obtained. One foot referred to stress on each syllable, two feet referred to a stressed syllable followed by an unstressed syllable, three feet referred to a stressed syllable, followed by two unstressed syllables and four feet to a stressed syllable, followed by three unstressed syllables and so on.

The stimuli preparation of rhythmic patterns was based on Fant's notion, where a ***"meter"*** is a sequence of recurrent feet in a regular pattern. Every metrical foot contains one strong syllable or beat and one or more weak syllable. Rhythm is a regular alternation between stressed (strong) and unstressed (weak) entities. In the current study, the relative strength factor of the syllable /ma/ was synthetically modified in three ways: viz, by altering the fundamental frequency or by altering intensity or by altering both fundamental frequency and intensity. Fundamental frequency was reduced in 5Hz steps and intensity was reduced in 10dB steps.

The following are the tokens synthesized and used in the verbal rhythm task under each condition (F0 manipulation, Intensity manipulation & F0-Intensity manipulation):

i. One foot condition
ii. Two feet condition
iii. Three feet condition
iv. Four feet condition
v. Five feet condition
vi. Six feet condition



(**Note**: - Stressed syllable & - Unstressed syllable)

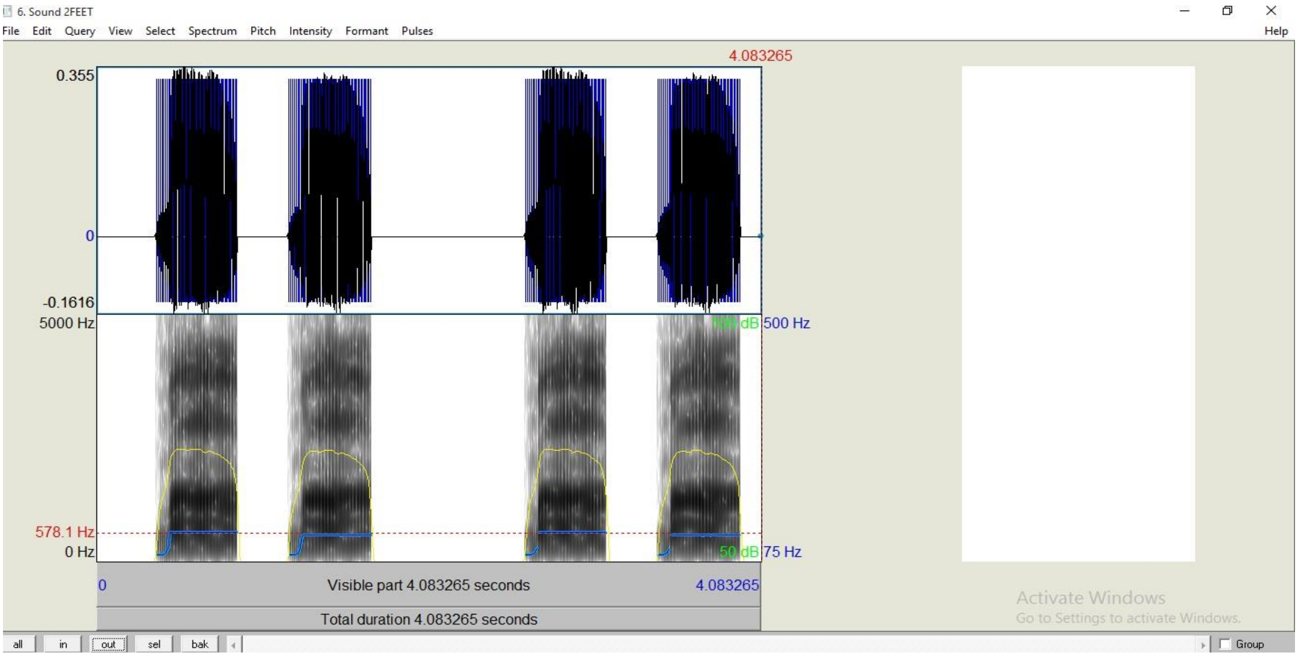


Figure 3.5. Screenshot of 2 feet stimuli in frequency manipulation condition (MA-ma-----MA-ma; MA: Stressed syllable & ma: Unstressed syllable)

One foot consisted of one strong (stressed) syllable /ma/, two feet comprised of one strong syllable followed by a weak (unstressed) syllable, etc. Six stimuli (1 foot, 2 feet, 3 feet, 4 feet, 5 feet & 6 feet) each in altered frequency condition, altered intensity and in altered frequency-intensity condition were synthesized. Thus a total of 18 stimuli (6 rhythm patterns X 3 manipulation conditions) were synthesized for the study.

b) Non-verbal rhythm task

For the non-verbal rhythm task, two sub-tasks modified from Corriveau and Goswami (2009) were used. The sub-tasks consisted of Simple (Paced) and Complex (Unpaced) conditions. Stimuli were created using online metronome (available on Google search page: Keyword search “online metronome”) with the following variations: 20 BPM (beats per minute) and 40 BPM. The task was computerized* and participants were asked to tap the response pad of the laptop (Lenovo G500 15.3inch) (the cursor was placed in the software’s

{*Computerized software is a stand-alone windows application developed using .net platform/ technology, C+ language and winamplibraries with the help of a qualified software engineer. The task was computerized to automatically record the time of tap by the participants to obtain Inter Tap Intervals (ITI)} response icon), in both paced (along with stimuli) and unpaced (without the stimuli) conditions.

In the unpaced condition, participants were instructed to remember the approximate inter-tap intervals (ITI) or the rhythm of the stimuli (based on previous paced condition) and tap the response pad of the laptop in similar pattern. The application was developed in order to automatically record the time of taps (in seconds) in each participant. Inter-tap interval (ITI) is the time period between successive taps in seconds and it can be computed based on the automatic output (time of taps in seconds) provided by the computerized software.

3.3.2. Phase II - Pilot study

i. Pilot study on adults

The stimuli tokens prepared were administered on five typical adults (SLPs) to obtain their views on the suitability of the stimuli tokens to test the prosody skills and the adequacy of the stimuli (in terms of inter-stimulus intervals and token durations). All 5 SLPs reached a conclusive decision to set 500 ms as stimulus duration and 1000 ms for the inter-stimulus interval (ISI). For the intonation task, the listeners were presented with various intonation patterns (High steady, Low steady, Rising and Falling) with a frequency range of 100 to 500 Hz. Stimuli levels above 400 Hz were perceived as very unnatural by the listeners. Hence the highest frequency of intonation patterns was set as 400 Hz. In the verbal rhythm task, listeners reported difficulty in imitation of 6 feet condition. The adult participants reported no difficulty in performing stress (verbal and non-verbal) and non-verbal rhythm task.

ii. Pilot study on children

Six typically developing (TD) children in the age range of 5-8 years participated in the pilot study. The following are the details of stimuli levels and responses of the participants.

a) Stress task (verbal & non-verbal)

- *Stimuli* - The synthetic tokens of syllable /ma/ was used as the stimuli for verbal stress task and tone glides were used for non-verbal stress task. The stimuli levels (25 stimuli pairs) of both tasks are listed in Table 3.1.

Table 3.1

Stimuli pairs for stress task (verbal and non-verbal)

Frequency manipulation tokens (in Hz)	**Intensity manipulation tokens (in dB)**	**Duration manipulation tokens (in sec)**
100-100	60-60	0.5-0.5
100-105	60-65	0.5-1
105-100	65-60	1-0.5
100-110	60-70	0.5-2
110-100	70-60	2-0.5
100-115	60-75	0.5-5
115-100	75-60	5-0.5
100-120	60-80	-
120-100	80-60	-

- *Response expected* - The participants were instructed to verbally indicate if the binary pair of stimuli presented in each token were perceived as same or different.

- *Observations* - Four participants in age range of 6-8 years required only 5 Hz difference to differentiate stressed token from unstressed, i.e, they were able to perceive and indicate the binary tokens to be different with the stimuli level of 100-105 Hz. However all the participants were able to identify 60 vs 70 dB level (i.e., 10 dB difference to perceive the stimuli as different).One 5.3 year old participant perceived difference in 100 vs 120Hz condition (that is, the participant required 20 Hz difference) and 60 vs 80 dB condition (needed 20 dB difference to perceive the tokens as different). All the participants were able to discriminate a 1-second difference in duration. (i.e., 0.5 vs 1.5 seconds stimuli).

b) Verbal intonation task

- *Stimuli*- The stimuli consisted of synthetically generated /a/ pitch glides. The range of frequency for the /a/ pitch glides was set as 100 Hz to 400 Hz (low frequency to high frequency). The frequency levels were decided based on the opinion from 5 SLPs as stated earlier (*refer page 51*).
- *Patterns* - High steady, Low steady, Rising and Falling
- *Response expected*-The participants were instructed to verbally imitate the perceived intonation pattern.
- *Observations*- All the six children were able to imitate the intonation patterns. Hence, the same glide range of 100-400 Hz was retained for the study.

c) Non-verbal intonation task

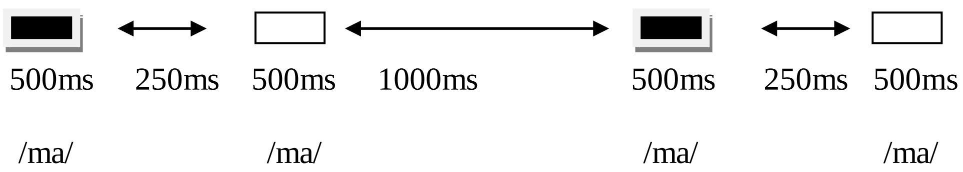
- *Stimuli*- The stimuli consisted of synthetically generated tone glides. The range of frequency for the tone glides was set as 100 Hz to 400 Hz (low frequency to high frequency).

- *Patterns*- High steady, Low steady, Rising and Falling
- *Response expected*- The participants were instructed to verbally imitate (hum) the perceived intonation pattern.
- *Observations*- All the six children were able to imitate the intonation patterns.

d) Verbal rhythm task

- *Stimuli*- The stimuli consisted of six rhythm patterns (with syllable /ma/) namely, 1 foot, 2 feet, 3 feet, 4 feet, 5 feet and 6 feet conditions.

  Each token of syllable /ma/ token was set to 500 ms and the inter-stimulus interval (between two sets of tokens) was set to 1000 ms. The inter-stimulus interval (between consecutive /ma/ syllables within one set of the token) was set to 250 ms.



(**Note:** ▬ - stressed syllable & ▭ - unstressed syllable)

The stressed /ma/ was synthetically modified to 120 Hz/ 80 dB (frequency/ intensity manipulation) and the lowest unstressed /ma/ was synthesized to 100 Hz/ 40 dB. That is, the frequency manipulation was in the range of 100 Hz (lowest unstressed syllable) to 120 Hz (highest stressed syllable) and intensity manipulation was in the range of 40-80 dB respectively. Fundamental frequency was reduced in 5Hz steps and intensity was reduced in 10 dB steps. There were a total of 3 manipulations in the tokens; Fundamental frequency (F0) only, Intensity only and F0-Intensity condition. Each manipulation consisted of the above mentioned 1 to 6 feet conditions. The tokens were initially presented to five typical adults and the participants reported difficulty with 6 feet condition. Hence, the 6 feet condition tokens were excluded from all the three manipulations and only tokens up to 5 feet condition were retained as the participants of the present study are children.

- *Response expected* - The children were instructed to verbally imitate the rhythmic patterns presented.
- *Observations* -All the children were able to imitate the tokens up to 5 feet rhythm condition. However, 2 children (one 5-year old and one from 7-year old) found the 5 feet rhythm pattern relatively difficult to imitate but were able to perform up to 4 feet condition.

e) Non-verbal rhythm task

- *Stimuli*- The stimuli consisted of recorded metronome variations with 20 Beats per Minute (BPM) and 40 BPM conditions. Each BPM consisted of 4 minute testing: 2 minutes paced condition (with stimuli) and 2 minutes testing for unpaced condition (without stimuli).
- *Response expected*- Participants were instructed to tap the response pad of the laptop (Lenovo G500 15.3inch), in accordance with the stimuli. The "Inter-tap intervals" were recorded automatically by the developed computerized application software (*refer page 10 for details*).
- *Observations*- All the six children of the pilot study preferred a 1-minute metronome stimuli, rather than 2-minute stimuli.

Based on the observations documented in the pilot study (on both adults and children), modifications were incorporated in terms of stimuli duration, inter-stimulus interval and stimuli levels.

3.3.3. Phase III-Administration of the developed tool on TDC and CwSLI

Procedure

Prior to the actual study, field testing was performed on five typically children in the age range of 5-8 years to ensure optimum stimuli levels in the school setting. Participants were made to sit comfortably on a chair, in a relatively quiet environment and the stimuli tokens were presented to them at comfortable listening level (75 dB SPL) through headphones (Philips Stereo SBC HP 100). The responses of the participants were audio recorded using Olympus WS 100 digital recorder.

Participants of Group I (typically developing children) and Group II (CwSLI) in the age range of 5-8 years were assessed using the verbal and non-verbal prosody tasks. Each participant was tested individually in a relatively quiet environment (Group I participants were tested in a relatively quiet room at school premises and Group II participants were tested in a relatively quiet room in the clinical setting). The duration of test administration was about 30 minutes for Group I participants and approximately 45 minutes for Group II participants. CwSLI (Group II) required more trials and repeated instructions during testing compared to TDC. The instructions for each task and procedure of test administration are provided in Appendix I (*refer to page no. 111*).

A total of **75 tokens** (25 stimuli pairs X 3 trials) **each for verbal and non-verbal stress tasks** were presented to the participants. For the intonation task, four patterns (that is, High steady, Low steady, Rising and Falling) were presented to all participants. A total of **12 intonation patterns** (4 patterns X 3 trials) **each for verbal and non-verbal intonation tasks** were presented. Administration of the **verbal rhythm task** consisted of a total of **45 stimuli tokens** (5 feet X 3 conditions X 3 trials). In **non-verbal rhythm task**, both **20 BPM and 40 BPM** condition was tested for **1-minute duration each** for all participants. The responses of all tasks were audio recorded except the non-verbal rhythm task, wherein the inter-tap intervals were automatically calculated by the developed software (*refer to page no.51) under Phase I for software details*.

3.4. Scoring

The verbal responses (same/different) of the stress task (verbal and non-verbal) were scored in terms of “0” and “1” by the experimenter. For the intonation (verbal and non-verbal) and rhythm (verbal) task, the responses of the participants and the respective stimuli tokens were appended together (eg: Stimuli-500ms -Response) with a 500 ms interval and saved as a single audio file for each token. This procedure was done for the ease of judges during perceptual analysis. The saved audio files were presented to three certified speech-language pathologists (SLPs) who served as judges for perceptual analysis. The three judges evaluated the imitations carefully listening to the audio tokens and indicate “Same/Different” depending on their perception. For the verbal rhythm task, two levels of scoring was done with ‘1’ score provided for the correct imitation of number of syllables in the stimuli token and ‘1’ additional score for the correct imitation of parameter (frequency/ intensity/ frequency-intensity) change in the stimuli token. Overall the criteria for a correct response included appropriate perception of stress, perception and usage of intonation and verbal rhythm (stimuli variations and number of syllables in the stimuli). The response “Same” was scored as “1” and the response “Different” were scored as “0” by the experimenter. For non-verbal rhythm task, mean and standard deviation of inter-tap intervals were computed for 20 BPM and 40 BPM conditions.

3.5. Reliability measures

3.5.1. Test-retest reliability

In the present study, for test re-test reliability, the investigator repeated the administration of the developed tool on 10 % of the participants i.e., 10 participants (including 9 TDC and 1 CwSLI) were tested again after two weeks from the initial administration. The test-retest reliability was found to be 90%.

3.5.2. Intra and Inter-judge reliability for the recorded responses

- For the intra-judge reliability the experimenter herself carefully listened to 10% of the responses again and scored the sample. Intra-judge reliability was found to be 100%.
- For inter-judge reliability, three certified SLPs participated as judges for analyzing 10 % of the data (refer 3.3. for details). The statistical measure of Kappa co-efficient was used to quantify inter-judge reliability and revealed moderate to good agreement.

3.6. Statistical analysis

The obtained data was analysed for normality and results revealed non-normal distribution ($p<0.05$). Hence, non-parametric tests such as Kruskal Wallis test and Mann Whitney U test were carried out.

CHAPTER IV

RESULTS AND DISCUSSION

The present study aimed to develop a test for assessment of prosody skills in children and establish norms for facilitating early identification of children with SLI. A total of 100 native Kannada speaking children in the age range of 5-8 years were considered as participants for the study. 90 typically developing children (TDC) were considered as Group I and 10 children with SLI (CwSLI) were considered as Group II. All the participants were instructed to perform prosody tasks targeting stress, intonation and rhythm. Results of the present study are discussed with respect to objectives of the study which are as follows:

1) To develop a test for assessing prosodic skills (stress, intonation & rhythm) in children.
2) To establish norms for prosodic skills for native Kannada speaking children using the developed test in the age range of 5-8 years.
3) To validate the developed test on 10 age-matched children with SLI (CwSLI).

Objective 1 - *To develop a test for assessing prosodic skills (stress, intonation & rhythm) in children.*

Based on the observations documented in the pilot study for the development of the test, modifications were incorporated in the stimuli tokens. After modifications, the stimuli tokens were administered on both Group I and Group II participants. Group I participants were able to perform in all three prosody tasks, namely, stress, intonation and rhythm with the modified stimuli tokens. Hence, the modified (post-pilot) tokens were considered as the stimulus in the finalized test. Table 4.1 describes the stimuli levels of the developed test.

Table 4.1

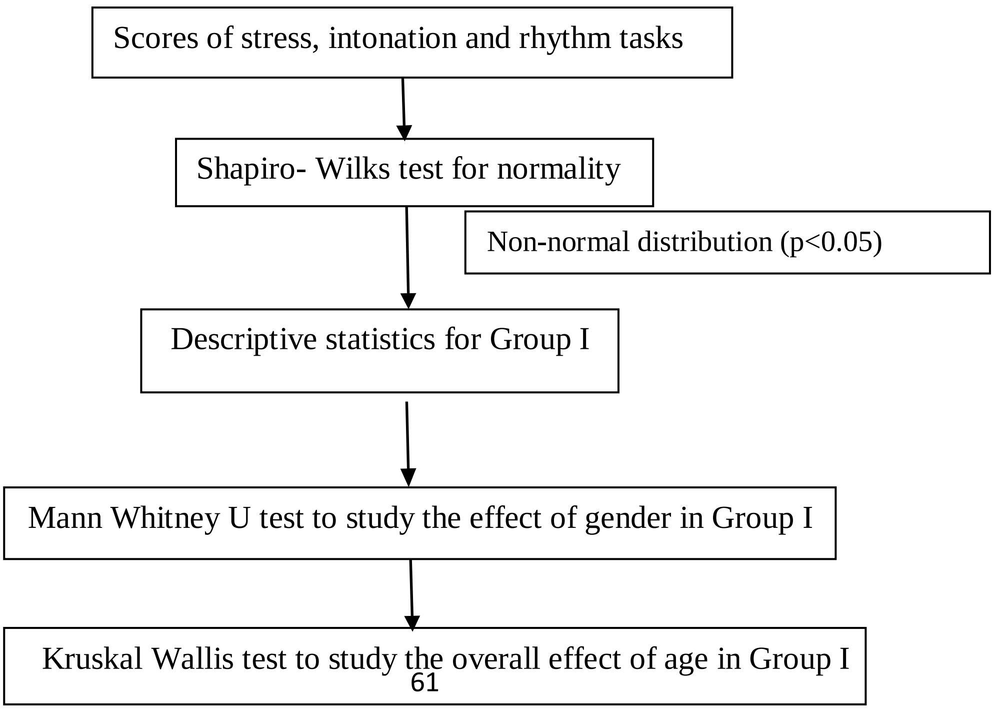
Stimuli details of the developed prosody test

Task	Verbal		Non-verbal	
	Stimuli Tokens	Stimuli levels	Stimuli Tokens	Stimuli levels
Stress	Syllable /ma/ 75 tokens (25 stimuli pairs X 3 trials)	Refer Table 3.1.	Piano tones 75 tokens (25 stimuli pairs X 3 trials)	Refer Table 3.1.
Intonation	/a/ pitch glides 12 tokens (4 patterns X 3 trials)	i. High steady (400Hz) ii. Falling (400Hz to 100 Hz) iii. Rising (100Hz to 400 Hz) iv. Low steady (100Hz)	Tone glides 12 tokens (4 patterns X 3 trials)	i. High steady (400Hz) ii. Falling (400 Hz to 100 Hz) iii. Rising (100Hz to 400 Hz) iv. Low steady (100Hz)
Rhythm	Syllable /ma/ 45 tokens (5 rhythm patterns X 3 conditions X 3 trials)	i.1 foot ii.2 feet iii.3 feet iv.4 feet v.5 feet	Recorded metronome audio files	i.20 Beats per Minute (BPM) ii.40 BPM

Objective 2-*To establish norms for prosodic skills for native Kannada speaking children using the developed test in the age range of 5-8 years.*

The developed test was administered on 90 typically developing children (TDC) in the age range of 5-8 years. In order to estimate test re-test reliability, the investigator repeated the administration of the developed tool on 10 % of the participants i.e., 10 participants (including 9 TDC and 1 CwSLI) were tested again after two weeks from the initial testing. Cronbach’s alpha test was carried out to compute the reliability coefficient for the test scores. A Cronbach’s alpha coefficient “α” of 0.7 and above is considered as statistically reliable. In the present study, the Crobach’s alpha for the scores were found to be 0.9, indicating good test re-test reliability.

The data was obtained in terms of accuracy scores for each task, which was subjected to test of normality. Results of Shapiro Wilks test revealed non- normal distribution (i.e., $p<0.05$). Hence, non-parametric tests such as Mann Whitney U test and Kruskal Wallis test were carried out for the statistical analysis of the data. The statistical analysis carried out for the data are summarized in Figure 4.1.



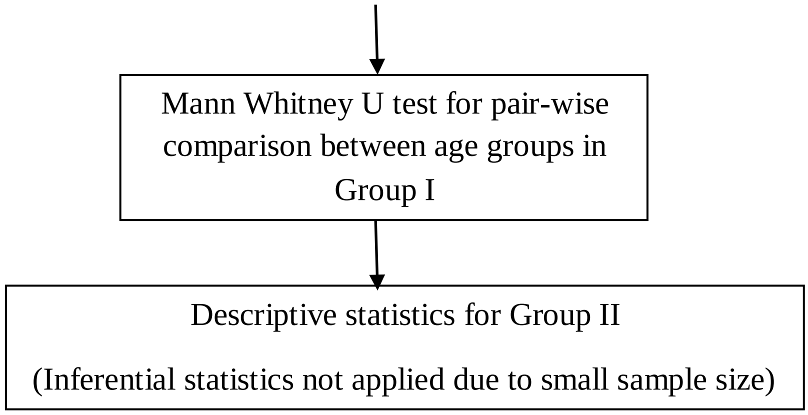


Figure 4.1. Flowchart depicting statistical analysis carried out for the current study

The descriptive statistics and effect of age and gender in prosody scores of Group I participants are presented for each task in the following section.

1. Stress task

A total of 75 stimuli tokens (*refer to Table 4.1*) each in verbal and non-verbal stress task were presented to the participants and responses in terms of accuracy scores were documented. The descriptive statistics of performance in verbal and non-verbal stress tasks are provided in Tables 4.2 and 4.3.

Table 4.2

Mean, Standard deviation and Median of performance scores in different age groups in verbal stress task in Group I (TDC)

Age groups	**Mean**	**Standard Deviation**	**Median**
5-6 years			
Boys	30.01	10.65	29.15
Girls	35.77	9.54	33.71
Average	30.90	10.86	29.57

6-7 years			
Boys	45.64	9.23	43.74
Girls	49.21	10.01	50.1
Average	47.44	9.60	48.12
7-8 years			
Boys	73.17	3.64	73.00
Girls	73.77	5.42	72.98
Average	73.47	4.57	71.52

Table 4.3

Mean, Standard deviation and Median of performance scores in different age groups in non- verbal stress task in Group I (TDC)

Age groups	**Mean**	**Standard Deviation**	**Median**
5-6 years			
Boys	65.13	8.52	61.54
Girls	66.89	8.40	65.77
Average	66.01	8.44	64.91
6-7 years			
Boys	71.75	4.59	70.83
Girls	73.53	3.61	72.15
Average	72.64	4.12	72.63
7-8 years			

Boys	74.21	2.86	72.10
Girls	74.87	0.02	73.46
Average	74.54	1.45	74.41

From Tables 4.2 and 4.3, it is observed that the mean values increased with age while the standard deviation decreased. The mean values of girls are higher than boys in all age groups.

It is also observed that the mean of non-verbal stress task are higher in all age groups when compared to verbal stress task.

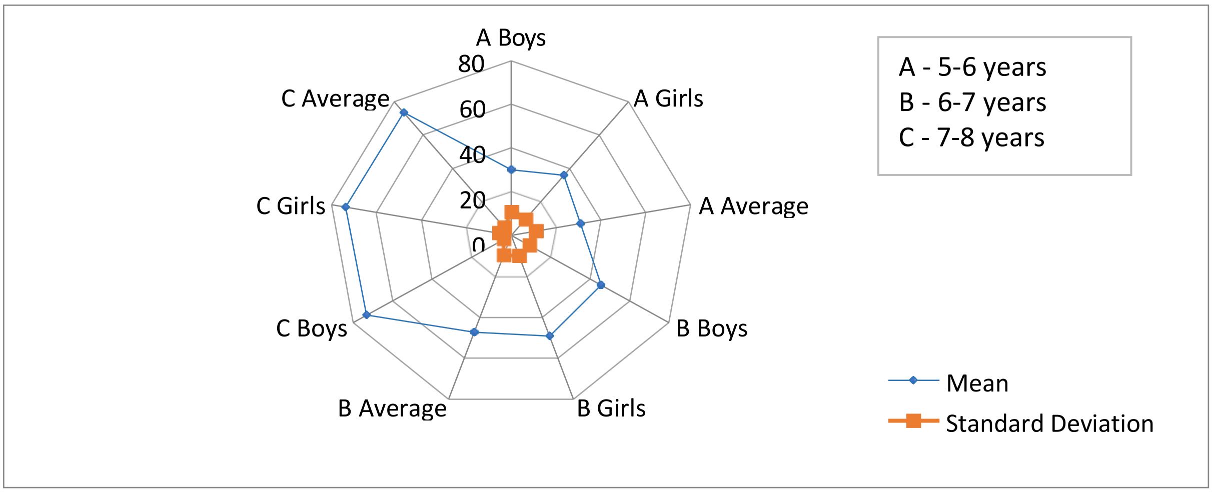


Figure 4.2. Mean and Standard deviation of performance scores in verbal stress task in Group I (TDC)

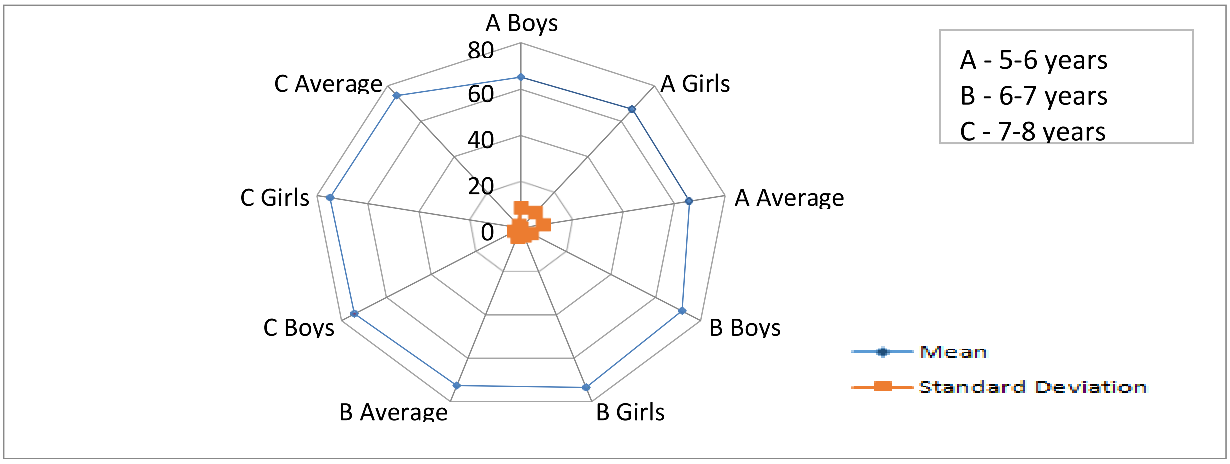


Figure 4.3. Mean and Standard deviation of performance scores in non-verbal stress task in Group I (TDC)

Figures 4.2.and 4.3 represent the mean and standard deviation of both verbal and non-verbal stress task in the form of radar graphs. The straight lines which divide the whole radar area into smaller parts are considered as the axes for the respective parts. Each of the axes is used to represent the degree of numerical scaling based on values of the data. More medial markings (towards the centre of radar) in the graph suggest lesser values and more lateral markings (away from the centre of radar) suggest greater values.

From Figure 4.2, it is inferred that there is a linear increase in mean with increase in age range (*observe the radar graph from right to left - clockwise direction*), as indicated by more lateral markings. The standard deviation decreases with increase in age and approach the value ‘0’ (centre of radar), indicating lesser variation in performance among the third age group (7-8 years) in verbal stress task.

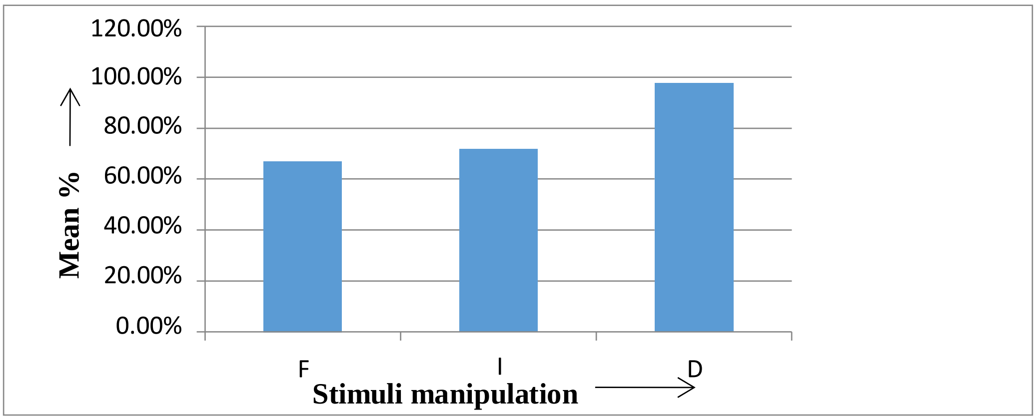
In non-verbal stress task (*refer to figure 4.3*) it can be noted that the distribution of markings are more lateral (i.e., mean values are greater) in all three age groups, indicating better performance in non-verbal than verbal stress task. Similar to the findings of verbal stress task, the graph also suggests a linear increase in mean and decrease in standard deviation with advancing age.

Mann Whitney U test was carried out to determine the overall effect of gender on stress perception and the results revealed *no gender effect* on stress perception in both verbal ($|z| = 0.06$; $p= 0.32$) and non-verbal ($|z| = 0.08$; $p= 0.27$) tasks across age groups. Results also suggested *no effect of gender within each age group.* Hence, the performance scores of both boys and girls were combined in each age group for further statistical analysis.

Kruskal Wallis test was carried out to compare the performance across all three age groups and results revealed an overall *significant difference ($p\leq0.05$)* in scores across age groups for both verbal ($\chi2$ =44.37; p= 0.02) and non-verbal ($\chi2$ =53.29; p= 0.02) stress tasks.

As Kruskal Wallis test revealed significant difference across age groups, Mann Whitney U test was carried out for pair-wise comparison between age groups and the results revealed significant pair-wise differences in performance between *5-6 years versus 6-7 years* (|z| = 4.23; p= 0.01), *6-7 years versus 7-8 years* (|z| = 5.35; p= 0.01) and *5-6 years versus 7-8 years* (|z| = 6.56; p=0.00). Overall, a *developmental trend for stress perception was observed* i.e., stress perception accuracy (mean scores) improved with age. This finding is in consonance with the observation reported by Kalathottukaren and Purdy (2017), suggesting an age-related improvement in prosody task performance especially in participants aged around 7 years. The study aimed to assess prosody perception in school-age children and concluded that much of the age-related changes occur in the age range of 7-8 years. In the current study participants in the age range of 7-8 years performed significantly better than those in 5-6 years and 6-7 years, signifying better prosody (*stress*) perception by 7 years. The results of the present study is also in agreement with the observation by Ballard, Djaja, Arciuli, James and Doorn (2012) who reported stress perception improves around 8 years of age and reaches the target (adult-like) by adolescence.

Mean percentage of scores were obtained from the available data for each stimuli manipulation condition (frequency/ intensity/ duration) to explore the major acoustic cue for stress perception. Mean percentage of scores according to each stimuli manipulation condition is represented in Figures 4.4 and 4.5.



(**Note**: F - Frequency manipulation, I - Intensity manipulation, D - Duration manipulation)

Figure 4.4. Mean percentage scores of verbal stress task in Group I (TDC)

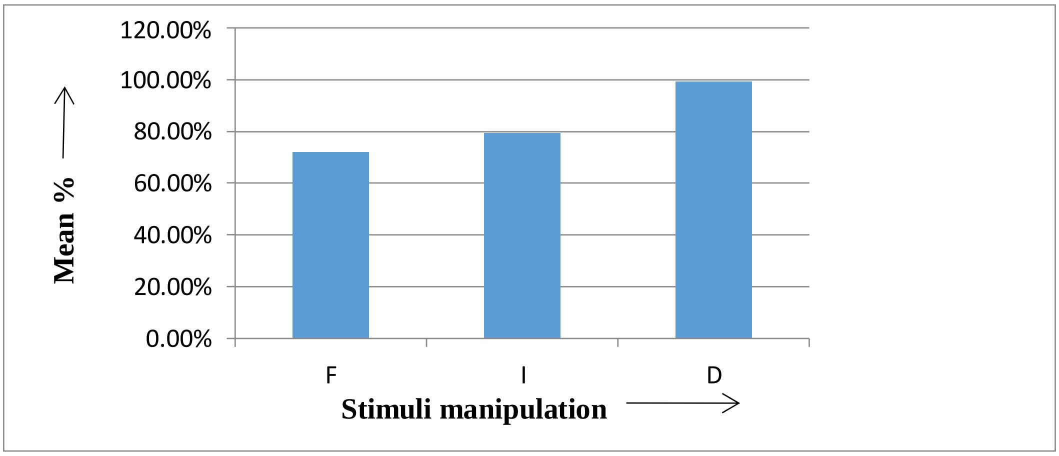


Figure 4.5. Mean percentage scores of non-verbal stress task in Group I (TDC)

Figures 4.4 and 4.5 indicate that the mean percentage scores are higher in duration manipulation condition followed by intensity and frequency manipulation conditions respectively. This finding is in agreement with the previous studies in Dravidian languages which state duration to be the major acoustic cue for stress perception. Similar observations have been noted in Kannada (Savithri, 1987; Rajupratap, 1991; Savithri, 1999a), Tamil (Balasubramanian, 1981) and Konkani (Kumar & Bhat, 2009) languages.

2. Intonation task

A total of 12 tokens (*refer to Table 4.1*) each for verbal and non-verbal intonation tasks were presented to all participants. The recorded imitation responses were analysed by three certified speech language pathologists (SLPs) (investigator - Judge 1 & other two SLPs - Judges 2& 3). The interjudge reliability measure was established for the intonation task scores of Group I participants. Kappa co-efficient was used to analyse the interjudge reliability and revealed moderate to very good agreement (0.40 to 1.0) (*refer to Appendix II*). The descriptive statistics of performance in verbal and non-verbal intonation tasks are provided in Tables 4.4 and 4.5 respectively. The maximum score for each intonation pattern was 3 (1 score X 3 trials= 3). Hence, calculating mean values for the data in each age group would result in values such as 2.61, 2.5, etc which would be a relatively insignificant statistical representation. Therefore, median and quartile deviation were calculated instead of mean and standard deviation.

Table 4.4

Median and Quartile deviation of performance scores in verbal intonation task for the three age groups in Group I (TDC)

Age Groups	**Intonation patterns**							
	High steady		**Falling**		**Rising**		**Low steady**	
	Median	**QD**	**Median**	**QD**	**Median**	**QD**	**Median**	**QD**
5-6 years								
Boys	3	0	2	1	3	0	3	0
Girls	2	1	1	1	3	0.5	3	0
Average	3	0.5	2	1.5	3	0.5	3	0

6-7 years								
Boys	3	0.5	2	1	3	0	3	0
Girls	3	0.5	2	0.5	3	0	3	0
Average	3	0.5	2	1.25	3	0	3	0
7-8 years								
Boys	3	0	2	0.5	3	0	3	0
Girls	3	0	3	0.5	3	0	3	0
Average	3	0	2.5	0.5	3	0	3	0

(**Note: QD**- Quartile Deviation)

Table 4.5

Median and Quartile deviation of performance scores in non-verbal intonation task for the three age groups in Group I (TDC)

Age groups	**Intonation patterns**							
	High steady		**Falling**		**Rising**		**Low steady**	
	Median	**QD**	**Median**	**QD**	**Median**	**QD**	**Median**	**QD**
5-6 years								
Boys	3	0.5	1	1	3	0	3	0
Girls	3	1	2	0.5	3	0	3	0
Average	3	0.5	2	0.5	3	0	3	0
6-7 years								
Boys	3	0	3	0.5	3	0	3	0
Girls	3	0	3	0.5	3	0	3	0
Average	3	0	3	0.5	3	0	3	0

7-8 years								
Boys	3	0	3	0	3	0	3	0
Girls	3	0	3	0.5	3	0	3	0
Average	3	0	3	0.5	3	0	3	0

From Tables 4.4 and 4.5, it is observed that the highest score of intonation task - '3'(score of '3' will be given if the participant had accurately imitated the perceived intonation pattern in all of the three trials) has been documented majorly in the *High steady, Rising and Low steady* intonation patterns in all age groups, which signifies that the *Falling* intonation pattern was comparatively more difficult to imitate for the participants. It is also observed that the median values are greater in the 7-8 year old age group, thereby suggesting an effect of age in imitating the *Falling* pattern. In terms of QD, it is found that with increase in age QD decreases signifying less variability in the higher age group. It can be noted that participants performed relatively better in non-verbal intonation task. By the age of 6-7 years, participants had high median and less QD in non-verbal intonation task while the same trend was achieved only by the age of 7-8 years in verbal intonation task. This finding suggests that participants were able to perceive all intonation patterns more accurately in case of non-verbal stimuli at a younger age.

The findings of the study are in concordance with the findings of Mohan Natarajan (1991). The author studied the imitation of intonation patterns among 4-8 year old children and concluded that with increase in age, the ability to imitate the intonation patterns increased. The present study's findings are also in consonance with Lai, Hughes and Shapiro (1991), who reported a developmental trend for affective processing (affective processing involves intonation perception). The finding of the present study is also in consonance with that of Koike and Asp (1981), who found effect of age on performance scores in task involving intonation patterns with the higher age group performing significantly better than lower age group.

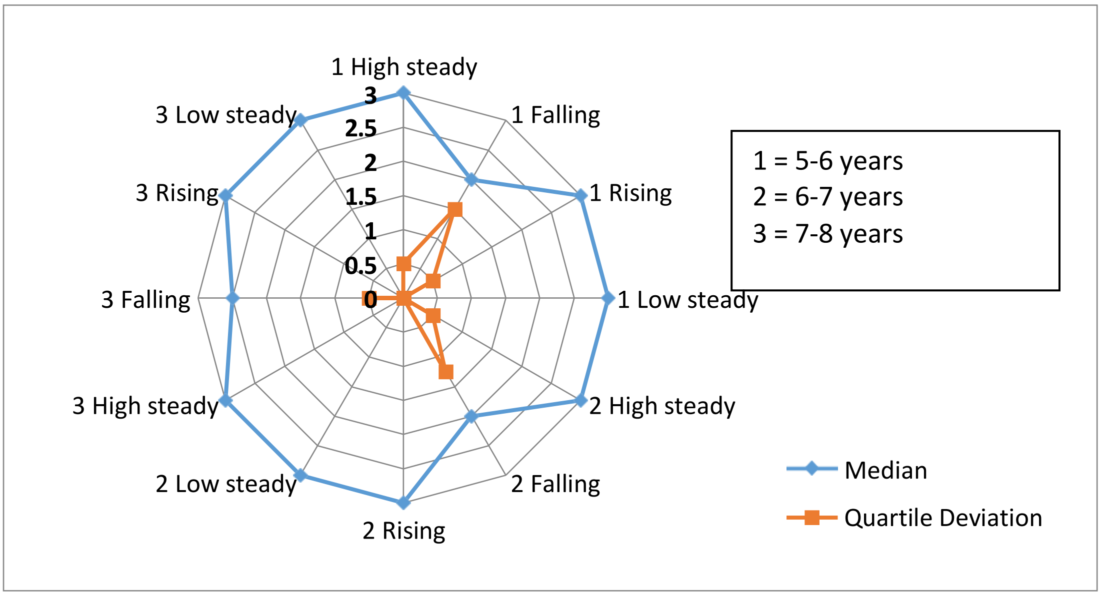


Figure 4.6. Median and Quartile deviation of intonation patterns in verbal intonation task in Group I (TDC)

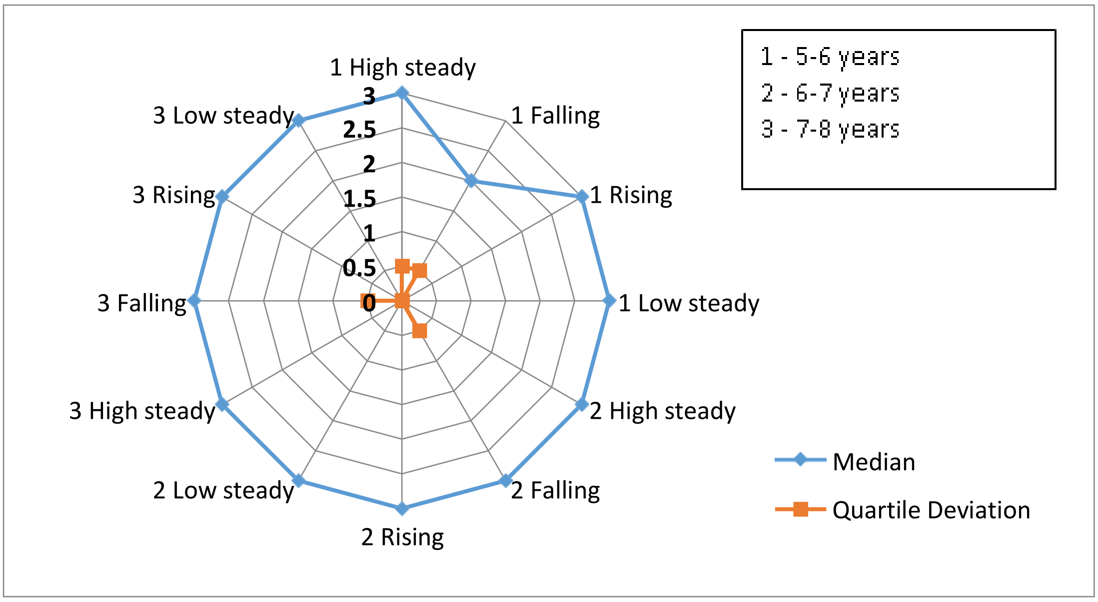


Figure 4.7. Median and Quartile deviation of intonation patterns in non-verbal intonation task in Group I (TDC)

Mann Whitney U test was carried out to determine the *effect of gender* on the scores of intonation task (verbal and non-verbal). The results revealed *significant difference* in *High* *steady* (|z|= 2.04; p=0.01) and *Falling* (|z|= 1.94; p=0.00) patterns between boys and girls in verbal intonation task in the *5-6 year age group* (where girls performed better than boys). No gender effect was observed in the 6-7 years and 7-8 year age groups for both verbal and non-verbal intonation tasks. This finding is different from the observation of Mohan Natarajan (1991), as the author reported no effect of gender in imitation of intonation patterns.

Kruskal Wallis test was carried out to compare the scores of intonation task (both verbal and non-verbal) across age groups. Significant differences were noted across all three age groups for specific intonation patterns, the results of which are summarized in Table 4.6.

Table 4.6

Results of Kruskal Wallis test comparing performance in intonation task (verbal and non-verbal) across age groups in Group I (TDC)

Intonation pattern	$\chi2$	p value*
High steady (verbal)	5.58	0.04
Falling (verbal)	8.91	0.01
High steady (non-verbal)	10.96	0.00
Falling (non-verbal)	36.75	0.00
Rising (non-verbal)	6.13	0.04

(* 0.05 level of significance)

It was also observed that there was a significant difference (p =0.01) between the total scores of verbal and non-verbal intonation task across all age groups, with participants performing better in non-verbal intonation task. Mann Whitney U test was performed for pair-wise comparison between age groups. A significant difference was observed for the intonation patterns listed in Table 4.7.

Table 4.7

Pair-wise comparison of performance in verbal and non-verbal intonation tasks between age groups in Group I (TDC)

Age groups	Intonation pattern	|z|	p value*
5-6 vs 6-7	High steady (non-verbal)	2.04	0.04
	Falling (non-verbal)	5.18	0.00
6-7 vs 7-8	No significant differences		
5-6 vs 7-8	Falling (verbal)	2.88	0.00
	High steady (non-verbal)	3.02	0.00
	Rising (non-verbal)	1.76	0.01
	Falling (non-verbal)	5.00	0.00

(* 0.05 level of significance)

No significant differences were observed when comparing the performance of participants in 6-7 and 7-8 years. However, in the results of Mann Whitney U test, the mean ranks of the 7-8 year group were relatively higher to that of 6-7 year group across all intonation patterns in both verbal and non-verbal intonation tasks.

3. Rhythm task

i. Verbal rhythm task

The recorded verbal imitation of the participants were analysed by two certified speech language pathologists (SLPs). The interjudge reliability measure was established for the administration of the verbal rhythm task on Group I participants. Kappa co-efficient was used to analyse the interjudge reliability which revealed a moderate to very good agreement (*refer* *to Appendix III*). Hence, the scores provided by the investigator (judge 1) were retained for further statistical analysis. The descriptive statistics of verbal rhythm task are detailed in Table 4.8.

Table 4.8

Mean, Standard deviation and Median of performance scores in verbal rhythm task in Group I (TDC)

Age groups	**Frequency manipulation (1 to 5 feet)**			**Intensity Manipulation (1 to 5 feet)**			**Frequency-Intensity manipulation (1 to 5 feet)**		
5-6 years	**M**	**SD**	**Med**	**M**	**SD**	**Med**	**M**	**SD**	**Med**
Boys	17.86	2.53	17.52	19.06	2.57	19.00	20.4	1.63	19.91
Girls	18.6	2.52	18.4	18.81	1.74	18.97	20.53	1.59	20.48
Average	18.23	2.52	18.35	18.93	2.15	18.76	20.46	1.61	20.24
6-7 years									
Boys	23.6	1.29	22.83	20.46	2.03	20.37	24.2	2.00	22.00
Girls	24.2	1.65	24.18	20.33	2.46	20.00	23.8	1.37	23.14
Average	23.9	1.47	23.72	20.39	2.24	20.13	24.00	1.68	22.89
7-8 years									
Boys	25.26	2.89	24.91	23.06	2.54	22.91	25.86	1.40	25.38
Girls	24.6	1.72	24.11	21.60	2.61	21.54	25.40	1.50	25.32
Average	24.93	2.30	24.99	22.33	2.57	22.57	25.63	1.45	25.61

(**Note: M**-Mean; **SD**-Standard Deviation; **Med**- Median; Mean scores are out of 30 in each age range: 5 feet conditions X 6 (total score for each token) = 30 - *Refer to method section for scoring details*)

From Table 4.8, it is inferred that participants performed better in combined frequency-intensity manipulation condition, followed by frequency alone and intensity alone manipulation conditions respectively. However, in 5-6 years age group, participants had higher mean scores in combined frequency-intensity manipulation condition followed by intensity alone and frequency alone conditions respectively. The mean values increased and SD values decreased linearly with increase in age.

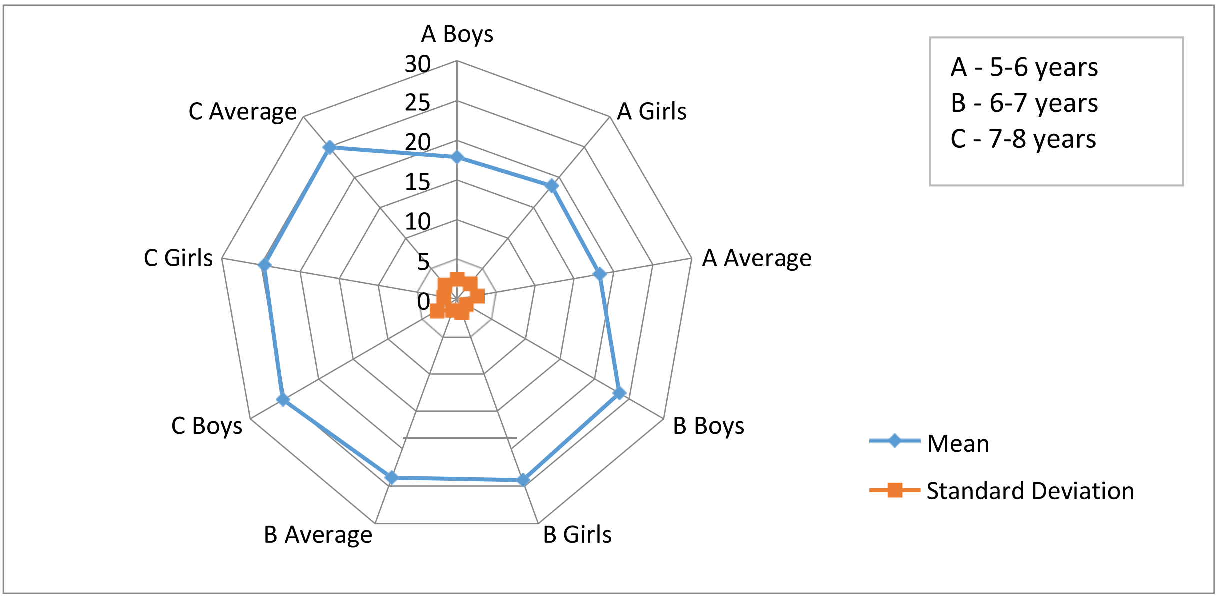


Figure 4.8. Mean and Standard deviation of frequency manipulation condition in verbal rhythm task in Group I (TDC)

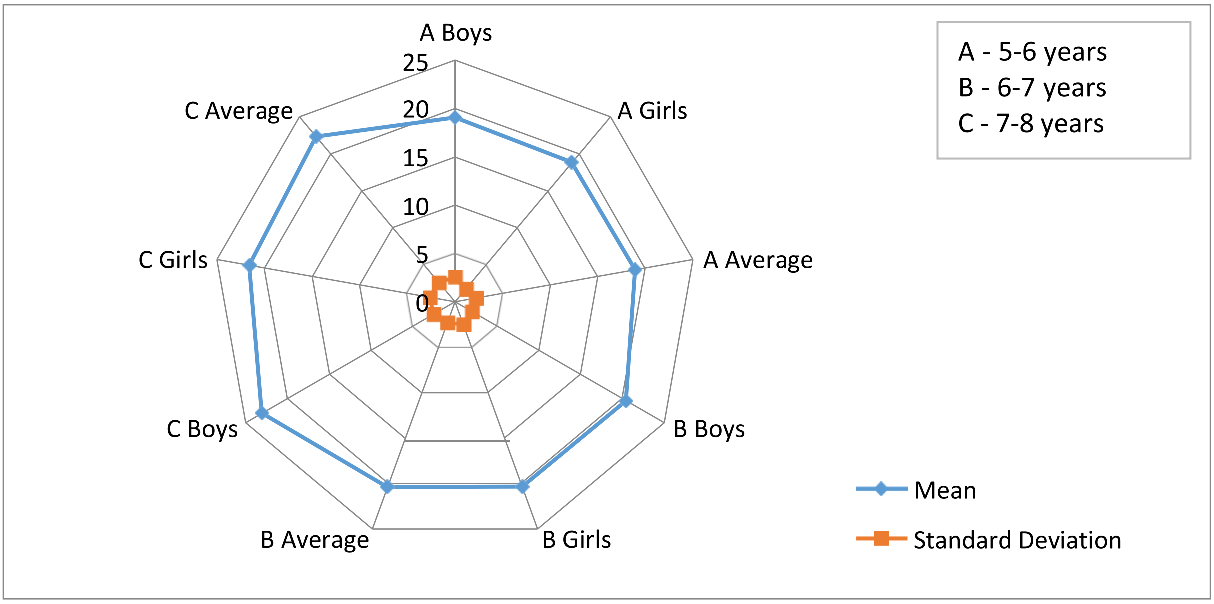


Figure 4.9. Mean and Standard deviation values of intensity manipulation condition in verbal rhythm task in Group I (TDC)

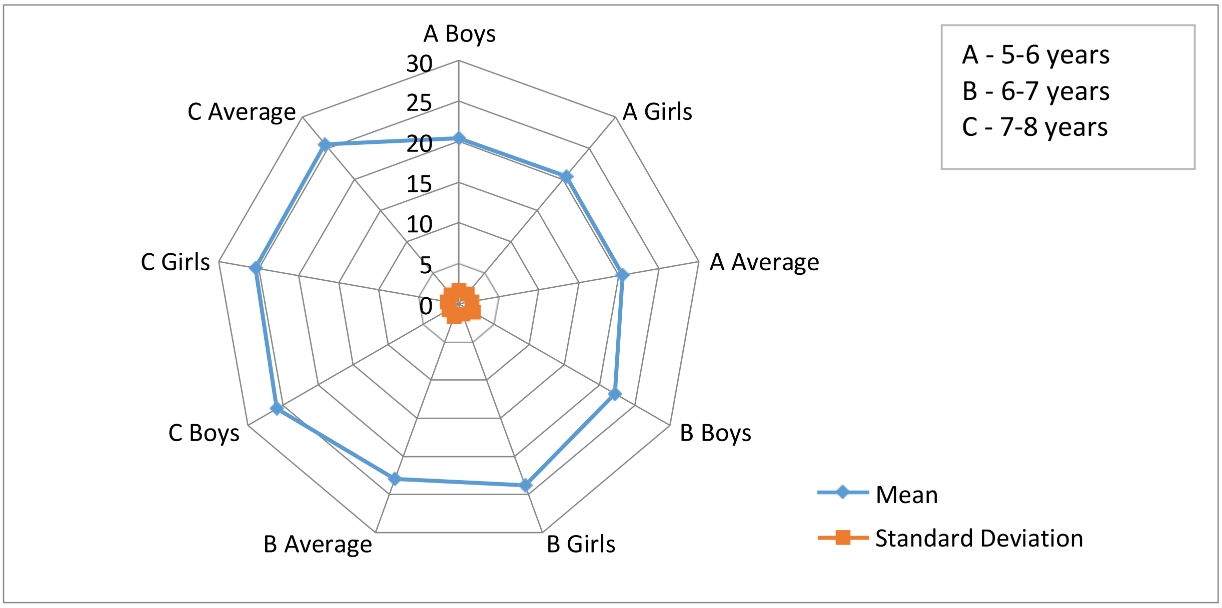


Figure 4.10. Mean and Standard deviation values of frequency-intensity manipulation condition in verbal rhythm task in Group I (TDC)

Table 4.9

Mean percentage scores of performance for frequency manipulation condition in verbal rhythm task in Group I (TDC)

Age range (in years)	**F-1**	**F-2**	**F-3**	**F-4**	**F-5**
5-6	100	71.6	68.33	31.17	10.52
6-7	100	100	97.23	63.88	37.28
7-8	100	100	94.4	72.75	45

(**Note**. F-1=1 foot frequency manipulation condition; F-2= 2 feet frequency manipulation condition and so on)

Table 4.10

Mean percentage scores of performance for intensity manipulation condition in verbal rhythm task in Group I (TDC)

Age range (in years)	**I-1**	**I-2**	**I-3**	**I-4**	**I-5**
5-6	100	100	72.22	32.29	11.13
6-7	100	100	73.41	40.55	26.1
7-8	100	100	76.67	55.68	40

(**Note**: I-1=1 foot intensity manipulation condition, I-2= 2 feet intensity manipulation condition and so on)

Table 4.11

Mean percentage scores of frequency-intensity manipulation condition in verbal rhythm task across age groups in Group I (TDC)

Age range (in years)	**FI-1**	**FI-2**	**FI-3**	**FI-4**	**FI-5**
5-6	100	99.44	89.44	39.23	12.2
6-7	100	100	97.77	66.15	34.6
7-8	100	100	98.33	82.77	45.5

(**Note**: FI-1=1 foot frequency-intensity manipulation condition, FI-2= 2 feet manipulation condition and so on)

From Tables 4.9, 4.10 and 4.11, it is observed that the mean percentage scores increase with increase in age. As the number of feet increases (from 1 foot to 5 feet), the mean percentage scores decrease. The results also indicate that the mean percentage scores are higher for combined frequency-intensity manipulation condition followed by frequency alone and intensity alone manipulation conditions respectively.

Mann Whitney U test was carried out to determine the effect of gender on scores of verbal rhythm task. The results revealed *significant difference in the total score of frequency-intensity manipulation* condition ($|z|=2.37$; $p=0.01$), wherein *girls performed better than boys in the 6-7 year age range*. *No gender differences* were observed *in the 5-6 year* and *7-8 year old* age groups. This finding is different from the observation of Jayanthi Ray (1993), who reported no gender differences in performance on rhythm task.

Kruskal Wallis test was carried out to compare scores of verbal rhythm task across age groups and significant differences noted are represented in Table 4.12.

Table 4.12

Results of Kruskal Wallis test comparing performance in verbal rhythm task across age groups in Group I (TDC)

Type of manipulation	Token	χ2	p value*
Frequency	F-2	9.845	0.00
	F-3	24.47	0.00
	F-4	21.98	0.00
	F-5	13.52	0.00
Intensity	I-3	10.8	0.00
	I-4	17.19	0.00
Frequency-Intensity	FI-3	11.38	0.00
	FI-4	29.60	0.01
	FI-5	24.39	0.00

(**Note:*** 0.05 level of significance; F-2 =2 feet frequency manipulation condition, 1-3 = 3 feet intensity manipulation condition, FI-3= 3 feet manipulation condition and so on)

The total scores of verbal rhythm task (χ2 =27.10; p=0.01) was also significantly different across all three age groups. Mann Whitney U test was carried out to determine pair-wise differences across age groups. The results of the same are summarized in Table 4.13.

Table 4.13

Pair-wise comparison of performance in verbal rhythm task between age groups in Group I (TDC)

Age range	Type of manipulation	Token	|z|	p value*
5-6 vs 6-7 years	**Frequency**	F-2	2.79	0.00
		F-3	3.89	0.00
		F-4	3.24	0.00
		F-5	2.50	0.01
	Frequency-Intensity	FI-4	3.65	0.00
		FI-5	2.29	0.02
6-7 vs 7-8 years	**Frequency**	F-4	1.14	0.01
	Intensity	I-2	1.91	0.05
		I-3	3.38	0.00
		I-4	3.38	0.00
	Frequency-Intensity	FI-3	2.18	0.02
		FI-4	3.42	0.00
		FI-5	4.25	0.00
5-6 vs 7-8 years	**Frequency**	F-3	4.32	0.00
		F-4	4.11	0.00
		F-5	3.55	0.00
	Intensity	I-3	2.20	0.02
		I-4	3.69	0.00
		I-5	2.97	0.00
	Frequency-Intensity	FI-3	3.33	0.00
		FI-4	4.63	0.00
		FI-5	4.37	0.00

(**Note:** * 0.05 level of significance; F-2 =2 feet frequency manipulation condition, 1-3 = 3 feet intensity manipulation condition, FI-3= 3 feet manipulation condition and so on)

The results also revealed that the total scores of verbal rhythm task (|z|=3.06; p=0.00) were significantly different between 5-6 year and 6-7 year age groups, with the latter group performing better. The participants in the age range of 7-8 years performed better than those in 6-7 years (|z|=3.71; p=0.00) and 5-6 years (|z|=4.38; p=0.00) in terms of the total scores of verbal rhythm task. Overall, it can be observed that with increase in age the mean percentage scores increased. The results of verbal rhythm task are supported by the observation of Jayanthi Ray (1993) in Kannada speaking children, who reported a similar linear developmental trend in the performance of rhythm task. The findings of the current study is in agreement with the observation of Koike and Asp (1981) who reported better performance in rhythm task with increase in age. In Kannada, studies have revealed speech rhythm patterns to be unclassified among children as they are in the developmental period of language acquisition (Savithri, Sreedevi & Kavya 2009 a. b). Literature suggests development of speech rhythm with increase in age and adult-like rhythm is likely to be achieved around 11-12 years (Savithri, Sreedevi, Jayakumar & Kavya, 2010). Hence, it can be inferred that with increase in age, rhythm perception and production improves. This statement supports the present study's finding as the scores of verbal rhythm task were found to increases with age.

ii. Non-verbal rhythm task

The non-verbal rhythm task consists of recorded metronome variations of 20 BPM (Beats Per Minute) and 40 BPM. Each BPM consist of paced (with stimuli) and unpaced conditions (without stimuli). The descriptive statistics of Group I participants in terms of inter-tap intervals (ITI; Time interval between two successive taps) are provided in Table 4.14.

Table 4.14

Mean, Standard deviation and Median of inter-tap intervals in non-verbal rhythm task in Group I (TDC)

Age range (years)	**P-20 (in sec)**			**P-40 (in sec)**			**UP-20 (in sec)**			**UP-40 (in sec)**		
	M	SD	Med	M	SD	Med	M	SD	Med	M	SD	Med
5-6												
Boys	3.64	0.30	3.58	1.76	0.31	1.74	4.19	0.60	4.23	2.26	0.52	2.19
Girls	3.67	0.27	3.65	2.02	0.19	2.15	4.26	0.42	4.28	2.29	0.33	2.24
Total	3.65	0.28	3.68	1.89	0.25	1.78	4.22	0.51	4.18	2.27	0.42	2.17
6-7												
Boys	2.44	0.33	2.42	1.91	0.21	1.87	2.59	0.33	2.48	2.25	0.27	2.18
Girls	2.49	0.41	2.38	2.08	0.25	2.11	3.09	0.65	3.00	2.04	0.35	1.97
Total	2.46	0.37	2.37	1.99	0.23	1.86	2.84	0.49	2.73	2.14	0.32	1.99
7-8												
Boys	2.90	0.31	2.94	1.65	0.14	1.69	3.14	0.48	3.20	1.76	0.24	1.74
Girls	3.20	0.15	3.19	1.65	0.17	1.50	3.62	0.52	3.58	1.99	0.37	1.86
Total	3.05	0.23	2.97	1.58	0.15	1.54	3.38	0.74	3.31	1.87	0.30	1.85

(**Note**: M- Mean; SD- Standard Deviation; Med- Median P-20 - Paced 20 Beats per minute condition; P-40 - Paced 40 BPM; UP-20 - Unpaced20 BPM; UP-40 -Unpaced 40 BPM)

The target mean ITI for 20 BPM and ITI for 40 BPM conditions are 3s and 1.5 s respectively. From Table 4.14, it can be observed that the mean ITI decreases as a function of age. Hence, it can be suggested that with increase in age, the mean ITI of the participants approximates towards the target ITI. This finding is in consonance with the observation reported by Corriveau and Goswami (2009), where the mean ITI of the typically developing participants were better (near to the target ITI) with less standard deviation. This trend was observed with increase in age.

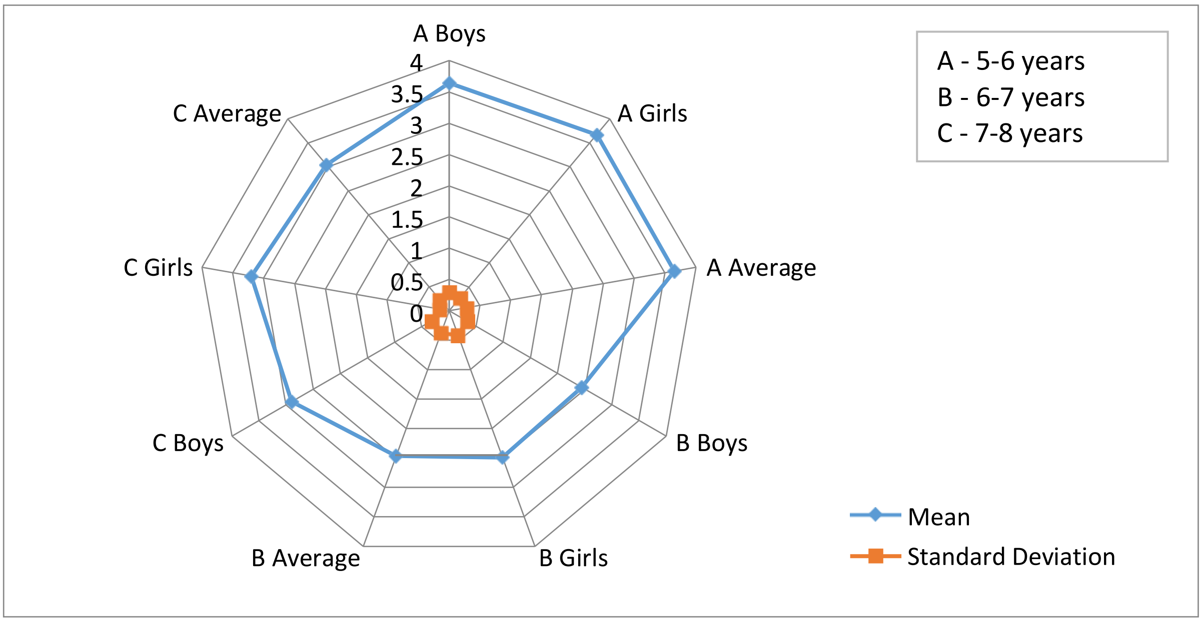


Figure 4.11. Mean and Standard deviation of inter-tap intervals in 20 BPM paced condition for non-verbal rhythm task in Group I (TDC)

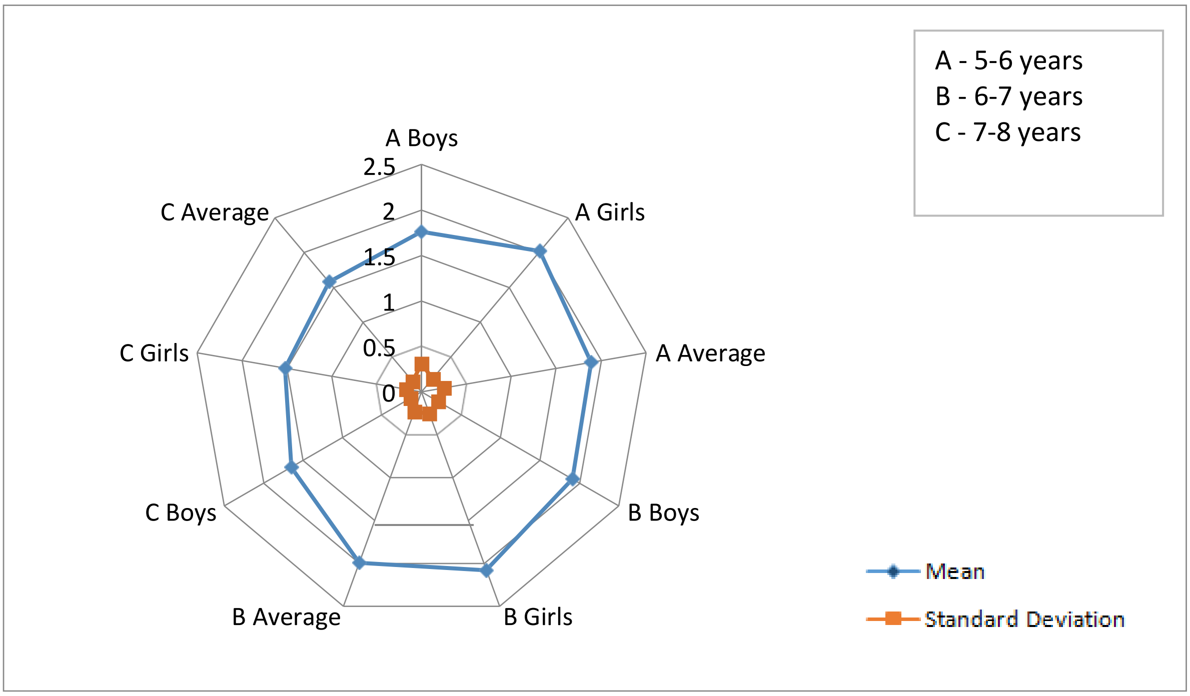


*Figure 4.12.*Mean and Standard deviation of inter-tap intervals in 40 BPM paced condition for non-verbal rhythm task in Group I (TDC)

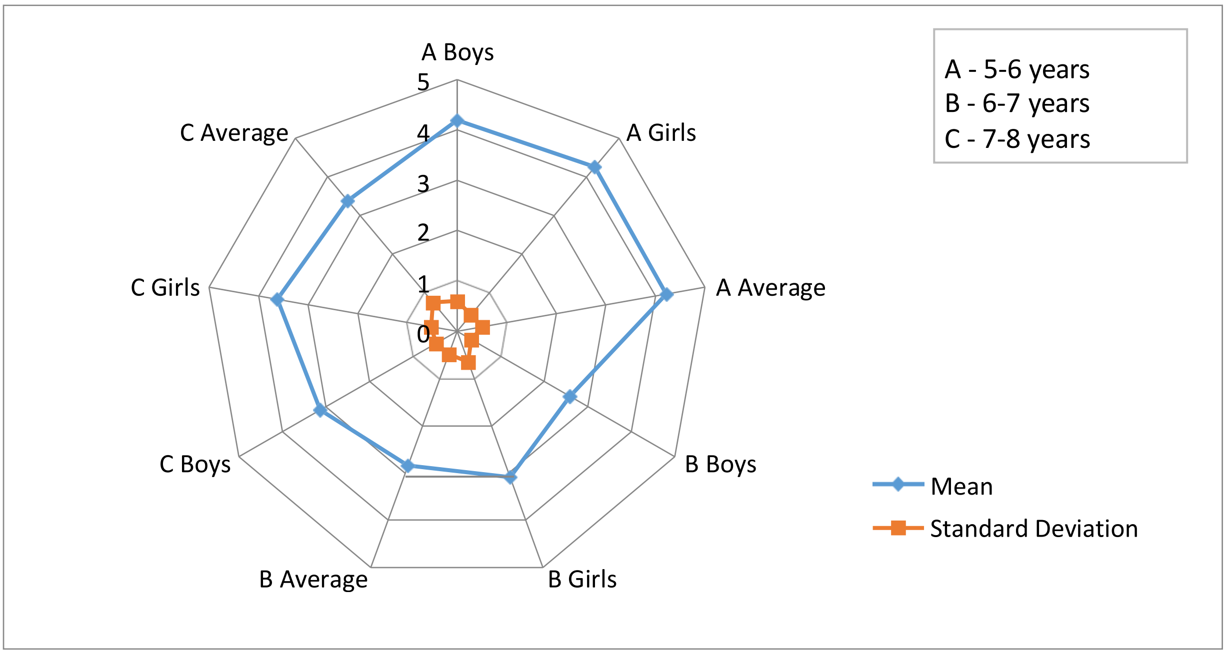


*Figure 4.13.*Mean and Standard deviation of inter-tap intervals in 20 BPM unpaced condition for non-verbal rhythm task in Group I (TDC)

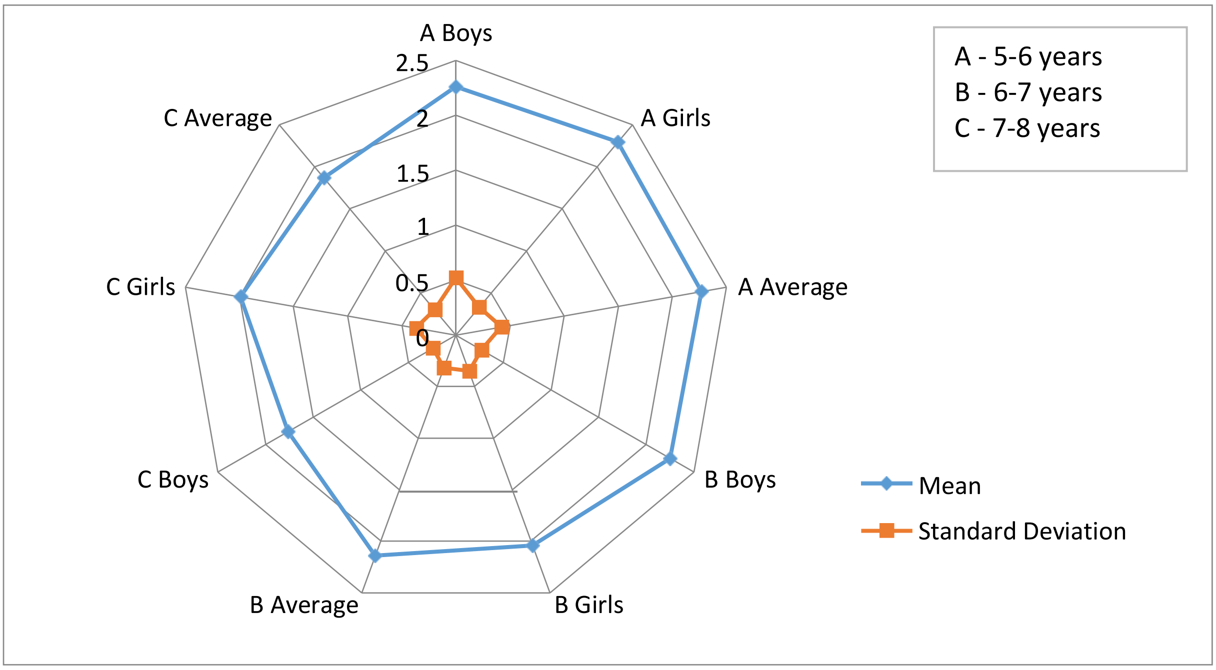


*Figure 4.14.*Mean and Standard deviation of inter-tap intervals in 40 BPM unpaced condition for non-verbal rhythm task in Group I (TDC)

Mann Whitney U test was carried out in order to determine the effect of gender. *Overall there was no gender effect* observed across age groups. Significant difference ($p \leq 0.05$) was observed in ITI-P20 ($|z|= 3.03$; $p= 0.02$), ITI-P40 ($|z|= 2.04$; $p=0.04$) and ITI-UP40 ($|z|= 2.45$; $p= 0.01$) conditions between males and females in 7-8 year old age group, wherein girls performed better than boys. No gender differences were noted in 5-6 and 6-7 year age groups. Kruskal Wallis test was carried out to compare the task performance across age groups and results revealed significant differences which are summarized in Table 4.15.

Table 4.15

Results of Kruskal Wallis test comparing performance in non-verbal rhythm task across age groups in Group I (TDC)

(* 0.05 level of significance)

Stimuli	**χ 2**	**p value***
P-20	65.95	0.00
P-40	34.41	0.00
UP-20	46.31	0.00
UP-40	15.85	0.00

From Table 4.15, it is observed that there was a significant difference in performance across the age ranges for all four stimuli conditions (P-20, P-40, UP-20 and UP-40).

Mann Whitney U test was performed for pair-wise age group comparisons, the results of which are summarized in Table 4.16.

Table 4.16

Pair-wise comparison of performance in non-verbal rhythm task across age groups in Group I (TDC)

Age groups (in years)	Stimuli	|z|	p value*
5-6 vs 6-7	P-40	6.62	0.00
	UP-40	5.91	0.00
6-7 vs 7-8	UP-40	5.42	0.00
5-6 vs 7-8	P-40	6.78	0.00
	UP-20	6.92	0.00
	UP-40	5.87	0.00

(* 0.05 level of significance)

Based on pair-wise comparison across age groups, it was observed that participants in the higher age group performed better than the other two groups. A similar developmental pattern in stress and intonation tasks was observed in the performance of rhythm task (verbal and non-verbal).

Objective 3 - *To validate the developed test on 10 age-matched children with SLI (CwSLI).*

The present study included 10 children (7 boys and 3 girls) with SLI (CwSLI) who constituted Group II. Due to reduced sample size and unequal age distribution of participants (n=3 in 5-6 years, 5 in 6-7 years and 2 in 7-8 years), statistical analysis to study the effect of variables such as age and gender was not possible. Hence, the following section describes the scores of Group II participants using basic essential statistical measures such as mean, standard deviation and median.

1. *Stress task*

The stimuli tokens (n=75; *refer to Table 4.1*) in each verbal and non-verbal stress task were presented to the participants and accuracy scores were noted. The descriptive statistics of performance in verbal and non-verbal stress tasks are provided in Table 4.17.

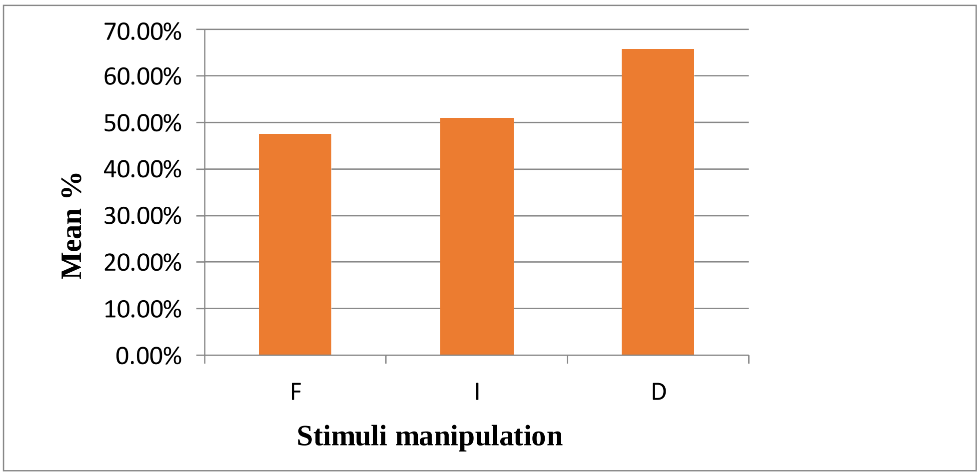
Table 4.17

Mean, Standard deviation and Median of performance scores of Group II (CwSLI) in verbal and non-verbal stress tasks

Stress task	Mean	SD	Median
Verbal	46.1	10.22	43.51
Non-Verbal	62.75	9.52	59.96

From Table 4.17, it is observed that the mean values are higher and standard deviation is lower for non-verbal stress task when compared to verbal stress task. A similar pattern with better performance in non-verbal stress task was followed by Group I participants (*refer to Tables 4.1 and 4.2*). However, CwSLI (Group II) had lower mean and higher standard deviation in both verbal and non-verbal stress tasks compared to TDC (Group I).

Mean percentage of scores were obtained from the available data for each stimuli manipulation condition (frequency/ intensity/ duration) to explore the major acoustic cue for stress perception. Mean percentage of scores according to each stimuli manipulation condition is represented in Figures 4.15 and 4.16.



(**Note**: F - Frequency manipulation, I - Intensity manipulation, D - Duration manipulation)

Figure 4.15. Mean percentage scores of verbal stress task in Group II (CwSLI)

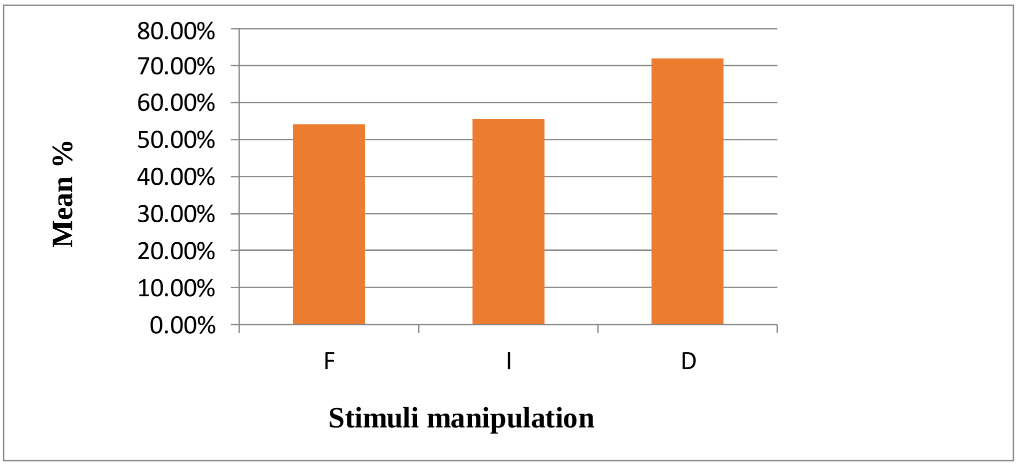


Figure 4.16. Mean percentage scores of non-verbal stress task in Group II (CwSLI)

Figures 4.15 and 4.16 indicate that the mean percentage scores are higher in duration manipulation condition followed by intensity and frequency manipulations respectively. Similar observation was found in Group I (TDC) participants (refer to Figure 4.4 and 4.5).This finding is in agreement with the previous studies in Dravidian languages which state duration to be the major acoustic cue for stress perception. Similar observations have been noted in Kannada (Savithri, 1987; Rajupratap, 1991; Savithri, 1999a), Tamil (Balasubramanian, 1981) and Konkani (Kumar & Bhat, 2009) languages.

2. *Intonation task*

The recorded imitation responses of the participants were analysed by three certified speech language pathologists (SLPs). Kappa co-efficient was used for inter judge reliability and results revealed moderate to very good agreement. The descriptive statistics of performance in verbal and non-verbal intonation tasks are provided in Table 4.18.

Table 4.18

Median and Quartile deviation of scores of Group II (CwSLI) in verbal and non-verbal intonation tasks

Intonation patterns	**Median**	**Quartile Deviation**
Verbal		
High steady	3	0.37

Falling	0	1.5
Rising	0	1.5
Low steady	3	0
Non-Verbal		
High steady	3	0
Falling	1.5	1.5
Rising	0	0.5
Low steady	3	0

From Table 4.18 it is observed that the highest score of intonation task (3) has been documented majorly in the steady patterns such as *High steady* and *Low steady* intonation patterns. It is also observed that changing intonation patterns such as *Falling* and *Rising* were more difficult for the participants to imitate. However, CwSLI (Group I) performed better in HL pattern in the non-verbal task when compared to verbal intonation task. A similar observation was documented with better performance in non-verbal intonation task in TD (Group I) participants also (refer to Tables 4.3 and 4.4). CwSLI had lower median and higher quartile deviation in both verbal and non-verbal intonation tasks compared to TDC.

3. *Rhythm task*

   i. Verbal rhythm task

The recorded responses of the participants were analysed by three certified speech language pathologists (SLPs). Kappa co-efficient was used for interjudge reliability and results revealed moderate to very good agreement. The descriptive statistics of performance in verbal rhythm task is provided in Table 4.19.

Table 4.19

Mean, Standard deviation and Median of performance scores of Group II (CwSLI) in verbal *rhythm task*

Stimuli manipulation	Mean	SD	Median
Frequency	15.3	2.54	15.1
Intensity	13.4	2.87	13.52
Frequency-Intensity	19.5	1.58	18.91

(**Note:** SD-Standard Deviation; Each stimuli manipulation consisted of 1 to 5 feet conditions)

From Table 4.19, it is inferred that participants performed better in combined frequency-intensity condition, followed by frequency alone and intensity alone manipulation conditions respectively. Similar findings in performance was observed in Group I (TDC) participants also (refer Table 4.7), however lower mean values and higher standard deviation was noted in Group II (CwSLI) participants.

i. Non-verbal rhythm task

The non-verbal rhythm task consists of recorded metronome variations of 20 BPM (Beats Per Minute) and 40 BPM. Each BPM consist of paced (with stimuli) and unpaced conditions (without stimuli). The descriptive statistics of Group II participants in terms of inter-tap intervals (ITI; time interval between two successive taps) are provided in Table 4.20.

Table 4.20

Mean, Standard Deviation and Median of inter-tap intervals of Group II (CwSLI) in non-verbal rhythm task

Stimuli condition	Mean	Standard Deviation	Median
P-20	4.21	0.47	4.06
P-40	1.96	0.26	1.94

UP-20	4.92	0.61	5.05
UP-40	2.54	0.43	2.58

(**Note**: P-20 - Paced 20 Beats per minute condition; P-40 - Paced 40 BPM; UP-20 - Unpaced20 BPM; UP-40 -Unpaced 40 BPM)

The target mean ITI for 20 BPM and 40 BPM conditions are 3s and 1.5s respectively. From Table 4.19, it is observed that the mean ITI of the participants approximates towards the target ITI in paced condition when compared to unpaced condition. Similar pattern was observed in Group I (TDC) participants also (refer to Table 4.14). However Group II (CwSLI) participants had higher mean ITI (than target ITI) when compared to mean ITI of Group I participants.

Based on the performance of Group II (CwSLI) participants in all three tasks (stress, intonation and rhythm) the following inferences are observed:

i. *Stress task* - Results revealed stress perception difficulties among CwSLI (indicated by lower mean and higher standard deviation values compared to TDC). This finding is supported by previous observations in children with language impairments (John & Howarth, 1965, Nickerson, 1975 and Graddol, 1991).

ii. *Intonation task* - Results revealed poor performance (lower mean and higher standard deviation) of CwSLI compared to TDC.The results of the current study are in agreement with all the previous studies (Clement et al, 2015; Sallat and Jenstchke, 2015; Akhila, 2016) wherein consistent evidence for deficit in intonation skills have been reported in CwSLI.

iii. *Rhythm task* - Results revealed reduced mean and increased standard deviation values in CwSLI compared to TDC in verbal rhythm task. In non-verbal rhythm task it was observed that Group I (TDC) participants performed better than Group II (CwSLI) in both paced and unpaced conditions. This finding is different from the observation by Corriveau and Goswami (2009) who reported that CwSLI were impaired in paced condition and not equally impaired (as in paced) in the unpaced condition. The difference in results can be attributed to the age of the participants, where they ranged from 7-11 years in the previous study and 5-8 years in the current study. Age related improvements in CwSLI might have lead to the observation that they are not equally impaired (like paced condition) in unpaced condition, as both TDC and CwSLI were reported to have similar level of difficulty in the previous study.

Intra-judge reliability for both Group I (TDC) and II (CwSLI) participants

The experimenter herself carefully listened to 10% of the participants responses again, that is, responses of 9 participants (10% of 90 participants) were randomly selected from Group I data and response of 1 participant (10% of 10 participants) was randomly selected from Group II data. The selected responses were scored again for intra-judge reliability and there was no observed difference between the test and re-test scores. Hence, the intra-judge reliability was noted to be 100% based on basic arithmetic calculation.

Highlights of results

- A developmental trend was observed in performance of all three tasks of prosody (stress, intonation and rhythm) among Group I (TDC) participants; with the 7-8 year age group performing significantly better.
- There was no overall effect of gender observed in any of the prosody tasks.
- Group I (TDC) participants performed significantly better than Group II (CwSLI) participants in all the prosody tasks and significant differences were noted.
- The mean/ median scores of TDC and CwSLI in stress, intonation and rhythm tasks are as follows:

1. Stress task (mean scores)

Task	Group I (TDC)			Group II (CwSLI)
	5-6 years	6-7 years	7-8 years	5-8 years
Verbal	30.90	47.44	73.47	46.10
Non-Verbal	66.01	72.64	74.54	62.75

2. Intonation task (median scores)

Task	Intonation patterns	Group I (TDC)			Group II (CwSLI)
		5-6 years	6-7 years	7-8 years	5-8 years
Verbal	HH	3	3	3	3
	HL	2	2	2.5	0
	LH	3	3	3	0
	LL	3	3	3	3
Non-Verbal	HH	3	3	3	3
	HL	2	3	3	1.5
	LH	3	3	3	0
	LL	3	3	3	3

3. Rhythm task

   a. Verbal rhythm task (mean scores)

Stimuli Manipulation	Group I (TDC)			Group II (CwSLI)
	5-6 years	6-7 years	7-8 years	5-8 years

Frequency	18.23	23.90	24.93	15.30
Intensity	18.93	20.39	22.33	13.40
Frequency-Intensity	20.46	24.00	25.63	19.50

Non-Verbal rhythm task

Stimuli Condition	**Group I (TDC)**			**Group II (CwSLI)**
	5-6 years	6-7 years	7-8 years	5-8 years
P-20	3.65	2.46	3.05	4.21
P-40	1.89	1.99	1.58	1.96
UP-20	4.22	2.84	3.38	4.92
UP-40	2.27	2.14	1.87	2.54

(**Note:** P-20= Paced 20 BPM; P-40= Paced 40 BPM; UP-20= Unpaced 20 BPM; UP-40= Unpaced 40 BPM)

4. All the participants (irrespective of age, gender and group) performed better in non- verbal subtasks than verbal subtasks of stress, intonation and rhythm.

CHAPTER V

SUMMARY AND CONCLUSIONS

The suprasegmental or prosodic patterns in a speech input convey significant cues with respect to lexical, syntactic and emotional aspects that help a listener comprehend the intended message of the speaker. The major suprasegmental aspects include stress, intonation, and rhythm. There are different functions of prosody namely, affective (conveys emotions and attitudes), grammatical (signals word boundaries) and pragmatic (context-based). Knowing how typically developing children perceive and process prosodic aspects during communication development is necessary. Literature suggests impaired prosody skills in children with communication disorders such as specific language impairment, hearing impairment, learning disability, autism spectrum disorder, etc,. In the Indian context, there are no published reports on norms for development of prosody skills, which can aid in identifying children at risk for communication disorders. Hence the current study aimed to develop a preliminary test for assessment of prosody in children and establish norms for facilitating early identification of SLI.

The study included a total of 100 native Kannada speaking children in the age range of 5-8 years. Group I consisted of 90 typically developing children (TDC) and group II consisted of 10 children with SLI (CwSLI). A pilot study was carried out among both adults and children to ensure optimum presentation of stimuli. The stimuli tokens were modified according to the documented pilot observations. Both verbal and non-verbal tasks of stress, intonation and rhythm were administered on all participants and the performance scores were noted. The scores of Group I participants were analyzed both with respect to age and gender in order to determine the significant differences, if any. A between-group {I (TDC) vs II (CwSLI)} comparison was carried out to determine the effect of group in the task scores. Data analysis revealed 100% intra-judge reliability and good inter-judge reliability (Kappa scores varied from *0.46 to 1.0*). Validity measures such as sensitivity and specificity were not determined as it requires a larger study population. Results revealed a developmental trend in performance of all three tasks (stress, intonation and rhythm) among Group I participants. There was no overall effect of gender observed in any of the prosody tasks. Group I participants performed significantly better than Group II participants in all the prosody tasks and significant differences were noted. All the participants (irrespective of age, gender and group) performed better in non-verbal tasks than verbal tasks of stress, intonation and rhythm.

Clinical implications

- The study provides insights to the developmental trend in perception and production of prosody across age in children. The findings of the study will serve as a normative for prosody skills among native Kannada speaking 5-8 year old children.
- The developed test is intended for clinical usage and the findings would aid in framing age-appropriate prosody goals during speech and language therapy of children with communication disorders.
- The results of the study suggest impaired prosody skills among children with SLI (CwSLI). The test developed can aid in identifying children at risk for Specific Language Impairment. Hence, prosody can be considered as one of the core areas to be focused in speech assessment and training of CwSLI.

Limitations of the study

- Intonation patterns used in the study are of basic types (High steady, Low steady, Rising and Falling). Other patterns such as gradual rise, steep rise, rise-fall, etc are not included.

- The small sample size and the unequal distribution of participants in the clinical group (age-wise) limits generalization of the results in this group.

Future directions

- Further studies can include participants with a wider age range to obtain more evidences on the developmental trajectory of each prosody dimension (stress, intonation and rhythm).
- The stimulus presentation and response formats can be computerized to make it user- friendly using suitable software programs for time efficient administration and to facilitate measurement of variables such as reaction time.
- The prosody test developed can be administered on a larger sample of CwSLI. This would help in setting a cut-off criterion in identifying children at risk for SLI.
- An experimental study involving comparison of language-age and chronological age matched CwSLI with typically developing children would strengthen the language- prosody link.
- The tasks can also be administered on other age-matched children with communication disorders such as learning disability, phonological disorder and hearing impairment to investigate the underlying prosody deficits in them. Stimuli duration and levels can be varied to explore minimum threshold levels of prosody perception in stress, intonation and rhythm tasks.
- The prosody test developed can be administered on participants with two etymologically different languages such as Indo Aryan (Hindi, Gujarathi, Bengali, etc) and Dravidian (Malayalam, Tamil, Telugu, etc) to explore how prosody skills vary across languages.

References

Abercrombie, D. (1964). *Syllable quantity and enclitics in English*. In honour of Daniel Jones, Abercrombie, D., Fry, D.F., McCarthy, P.A.D., Scott, N. C. & Trim, J.M.L., eds. London: Longmans.

Ainsworth, W. A. (1972). Duration as a cue in the recognition of synthetic vowels. *The Journal of the Acoustical society of America*, *51*(2B), 648-651.

Akhila (2016). Perception of affective prosody in 5-7 year old Malayalam speaking children with SLI. *Unpublished Masters Dissertation, University of Mysore.*

Allen, G. D. (1975). Speech rhythms: Its relation to performance universals and articulatory timing. *Journal of phonetics*, 3 (2), 75-86.

Asher, R. E. (1985). Tamil (Croom Helm Descriptive Grammars). *London: Croom Helm*.

Atkinson-King, K. 1973.Children's acquisition of Phonological stress contrasts. Papers in Phonetics, 25, Phonetics Laboratory, University of California, L.A.

Balasubramanian T. (1981). Emphasis in Tamil. *Journal of Phonetics*, 9(2),139-50.

Ballard, K. J., Djaja, D., Arciuli, J., James, D. G., & van Doorn, J. (2012). Developmental trajectory for production of prosody: lexical stress contrastivity in children ages 3 to 7 years and in adults. *Journal of Speech, Language, and Hearing Research*, 55 (6), 1822-1835.

Baltaxe, C. A. (1984). Use of contrastive stress in normal, aphasic, and autistic children. *Journal of Speech, Language, and Hearing Research*, *27*(1), 97-105.

Balusu, R. (2001). Acoustic correlates of stress and accent in Telugu. *21st South Asian Languages Analysis Roundtable, University of Konstanz*, 7-19.

Bertinnito, P. M. (1980). Perception of stress by Italians. *Journal of Phonetics*, *8*, 385-395.

Borrie, S. A., Lansford, K. L., & Barrett, T. S. (2017). Rhythm perception and its role in perception and learning of dysrhythmic speech. *Journal of Speech, Language, and Hearing Research*, *60*(3), 561-570.

Bowerman, M. (1982). 11. Reorganizational processes in lexical and syntactic development. *Language acquisition: The state of the art*, *319*.

Bowers, D., Blonder, L. X., & Heilman, K. M. (1999). The Florida Affect Battery. Gainesville, FL: University of Florida Brain Institute.

Brandt, A. K., Slevc, R., & Gebrian, M. (2012). Music and early language acquisition. *Frontiers in psychology*, *3*, 327.

Breen, M., Kaswer, L., Van Dyke, J. A., Krivokapić, J., & Landi, N. (2016). Imitated prosodic fluency predicts reading comprehension ability in good and poor high school readers. *Frontiers in Psychology*, *7*, 1026.

Choi, J. Y., Hasegawa-Johnson, M., & Cole, J. (2005). Finding intonational boundaries using acoustic cues related to the voice source. *The Journal of the Acoustical Society of America*, *118*(4), 2579-2587.

Clément, S., Planchou, C., Béland, R., Motte, J., & Samson, S. (2015). Singing abilities in children with Specific Language Impairment (SLI). *Frontiers in psychology*, *6*, 420.

Collier, R. P. G. (1991). Multi-language intonation synthesis. *Journal of Phonetics*, *19*, 61-73.

Collins, B. & Mees M.I. (2003). Practical Phonetics and Phonology, A resource book for students, Routledge, pp. 109-116

Corriveau, K. H., & Goswami, U. (2009). Rhythmic motor entrainment in children with speech and language impairments: tapping to the beat. *cortex*, *45*(1), 119-130.

Crary, M. A., & Tallman, V. L. (1993). Production of linguistic prosody by normal and speech-disordered children. *Journal of Communication Disorders*, *26*(4), 245-262.

Crystal, D. (1973). Non-segmental phonology in language acquisition: A review of the issues. *Lingua*, *32*(1-2), 1-45.

Crystal, D. (1982). Profiling linguistic disability. London, UK: Edward Arnold.

Cumming, R. E. (2010). The interdependence of tonal and durational cues in the perception of rhythmic groups. *Phonetica*, *67*(4), 219-242.

Cumming, R., Wilson, A., Leong, V., Colling, L. J., & Goswami, U. (2015). Awareness of rhythm patterns in speech and music in children with specific language impairments. *Frontiers in human neuroscience*, *9*, 672.

Cutler, A., & Butterfield, S. (1992). Rhythmic cues to speech segmentation: Evidence from juncture misperception. *Journal of memory and language*, *31*(2), 218-236.

Cutler, A., & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human perception and performance*, *14*(1), 113.

Darley, F. L., Aronson, A. E., & Brown, J. R. (1969b). Clusters of deviant speech dimensions in the dysarthrias. *Journal of speech and hearing research*, *12*(3), 462-496.

dePijper, J. R., & Sanderman, A. A. (1994). On the perceptual strength of prosodic boundaries and its relation to suprasegmental cues. *The Journal of the Acoustical Society of America*, *96*(4), 2037-2047.

Divya (2005). Manual for correction of intonation in Kannada speaking children with speech and language disorders. *Unpublished Masters Dissertation, University of Mysore.*

Fisher, J., Plante, E., Vance, R., Gerken, L., & Glattke, T. J. (2007). Do children and adults with language impairment recognize prosodic cues?. *Journal of Speech, Language, and Hearing Research*.

Fuchs, R. (2016). Speech Rhythm in Varieties of English. In *Speech Rhythm in Varieties of English* (pp. 87-102). Springer, Berlin, Heidelberg.

Geeta, S., & Muralidhara, B. L. (2017, March). Syllable as the basic unit for Kannada speech synthesis. In *2017 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)* (pp. 1205-1208). IEEE.

Gelfer, C. E. (1980). The ability to vary stress in a subject with apraxia of speech. *The Journal of the Acoustical Society of America*, *67*(S1), S79-S79.

Goswami, U. (2010). A temporal sampling framework for developmental dyslexia. *Trends in cognitive sciences*, *15*(1), 3-10.

Goswami, U., Gerson, D., & Astruc, L. (2010). Amplitude envelope perception, phonology and prosodic sensitivity in children with developmental dyslexia. *Reading and Writing*, *23*(8), 995-1019.

Guasti, M. T. (2017). *Language acquisition: The growth of grammar*. MIT press.

Hemakumar, G., & Punitha, P. (2013). Speech recognition technology: a survey on Indian languages. *International Journal of Information Science and Intelligent System*, *2*(4), 1-38.

Highnam, C., & Morris, V. (1987). Linguistic stress judgments of language learning disabled students. *Journal of communication disorders*, *20*(2), 93-103.

Holliman, A. J., Wood, C., & Sheehy, K. (2012). A cross-sectional study of prosodic sensitivity and reading difficulties. *Journal of Research in Reading*, *35*(1), 32-48.

Hood, R. B., & Dixon, R. F. (1969). Physical characteristics of speech rhythm of deaf and normal-hearing speakers. *Journal of Communication Disorders*, *2*(1), 20-28.

Hudgins, C. V. (1946). Speech breathing and speech intelligibility. *Volta review*, *48*(11), 642-644.

Hyde, K. L., & Peretz, I. (2004). Brains that are out of tune but in time. *Psychological Science*, *15*(5), 356-360.

Ito, K. (2014). Children's pragmatic use of prosodic prominence. *Pragmatic development in first language acquisition*, *10*, 199-218.

Järvinen-Pasley, A., Peppé, S., King-Smith, G., & Heaton, P. (2008). The relationship between form and function level receptive prosodic abilities in autism. *Journal of Autism and Developmental Disorders*, *38*(7), 1328-1340.

Jassem, W., Morton, J., & Steffen-Batog, M. (1968). The perception of stress in synthetic speech-like stimuli by Polish listeners. *Speech analysis and synthesis*, *1*, 289-308.

Jayanthi Ray (1993). Synthetic test of rhythm. *Unpublished Masters Dissertation, University of Mysore.*

Jentschke, S., & Koelsch, S. (2009). Musical training modulates the development of syntax processing in children. *Neuroimage*, *47*(2), 735-744.

Jiménez-Fernández, G., Gutiérrez-Palma, N., & Defior, S. (2015). Impaired stress awareness in Spanish children with developmental dyslexia. *Research in Developmental Disabilities*, *37*, 152-161.

John, J. E. J., & Howarth, J. N. (1965). The effect of time distortions on the intelligibility of deaf children's speech. *Language and Speech*, *8*(2), 127-134.

Jusczyk, P. W., & Luce, P. A. (2002). Speech perception and spoken word recognition: Past and present. *Ear and hearing*, *23*(1), 2-40.

Kalathottukaren, R. T., Purdy, S. C., & Ballard, E. (2015). Behavioral measures to evaluate prosodic skills: A review of assessment tools for children and adults. *Contemporary Issues in Communication Science and Disorders*, *42*, 138.

Keane, E. (2006). Prominence in Tamil. *Journal of the International Phonetic Association*, *36*(1), 1-20.

Kenworthy, J. (1987). Teaching English pronunciation. New York: Longman

Kidd, G. R., Watson, C. S., & Gygi, B. (2007). Individual differences in auditory abilities. *The Journal of the Acoustical Society of America*, *122*(1), 418-435.

Killion, M. C., Niquette, P. A., Gudmundsen, G. I., Revit, L. J., & Banerjee, S. (2004). Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *116*(4), 2395-2405.

Klieve, S. A. (1998). Perception of prosodic features by children with cochlear implants. Is it sufficient for understanding meaning differences in language (Unpublished master's thesis). University of Melbourne, Australia.

Koelsch, S. (2011). Toward a neural basis of music perception–a review and updated model. *Frontiers in psychology*, *2*, 110.

Koike, K. J., & Asp, C. W. (1981). Tennessee test of rhythm and intonation patterns. *Journal of Speech and Hearing Disorders*, *46*(1), 81-87.

Kumar RB & Bhat JS. (2009). Acoustic correlates of stress in Konkani language. *Languages in India*. 9(4). 52-54.

Lai, Z., Hughes, S., & Shapiro, E. (1991). Manual for the Minnesota Tests of Affective Processing (MNTAP). *University of Minnesota, MN*.

Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology: Human perception and performance*, *21*(3), 451.

Lavie, N. (2000). Distracted and confused?: Selective attention under load. *Trends in cognitive sciences*, *9*(2), 75-82.

Lavie, N., Hirst, A., De Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General*, *133*(3), 339.

Lehiste, I. (1968a). Vowel quantity in word and utterance in Estonian. *Congress ussecundus international lisfinno-ugristarum, Helsinki 1965*, 293-303.

Lehiste, I. (1970). Suprasegmentals. Oxford University Press, Cambridge.

Lehiste, I. (1977) Isochrony reconsidered. Journal of Phonetics, 5, pp. 253-263.

Lochrin, M., Arciuli, J., & Sharma, M. (2015). Assessing the relationship between prosody and reading outcomes in children using the PEPS-C. *Scientific Studies of Reading*, *19*(1), 72-85.

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

Low, E. L., Grabe, E. & Nolan, F. (2000). Quantitative Characterizations of Speech Rhythm – 'Syllable timing' in Singapore English. *Language and Speech*, 43, 377 – 401.

Magne, C., Schön, D., & Besson, M. (2006). Musician children detect pitch violations in both music and language better than nonmusician children: behavioral and electrophysiological approaches. *Journal of cognitive neuroscience*, *18*(2), 199-211.

Manjula (1979). Intonation in Kannada: Some Aspects. *Unpublished Masters Dissertation, University of Mysore.*

Manjula (1997). Aspects of Kannada Intonation in discourse (Interrogatives). *Unpublished Masters Dissertation, University of Mysore.*

Mari, G., Scorpecci, A., Reali, L., & D'Alatri, L. (2016). Music identification skills of children with specific language impairment. *International journal of language & communication disorders*, *51*(2), 203-211.

Marshall, C. R., Harcourt-Brown, S., Ramus, F., & Van Der Lely, H. K. J. (2009). The link between prosody and language skills in children with specific language impairment (SLI) and/or dyslexia. *International Journal of Language & Communication Disorders*, *44*(4), 466-488.

Marshall, C. R., Harcourt-Brown, S., Ramus, F., & Van Der Lely, H. K. J. (2009). The link between prosody and language skills in children with specific language impairment (SLI) and/or dyslexia. *International Journal of Language & Communication Disorders*, *44*(4), 466-488.

Martin, J. G. (1979). Rhythmic and segmental perception are not independent. *The Journal of the Acoustical Society of America*, *65*(5), 1286-1297.

Mattys, S. L., White, L., & Melhorn, J. F. (2005). Integration of multiple speech segmentation cues: a hierarchical framework. *Journal of Experimental Psychology: General*, *134*(4), 477.

McQueen, J. M., Norris, D., & Cutler, A. (1994). Competition in spoken word recognition: Spotting words in other words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(3), 621.

Menyuk, P. (1972). Sentences children use. *MIT research monograph*, 52, 2-7.

Mok, P., &Dellwo, V. (2008, May). Comparing native and non-native speech rhythm using acoustic rhythmic measures: Cantonese, Beijing Mandarin and English. In *Proceedings of Speech Prosody* (Vol. 4, pp. 423-426).

Monsen, R. B. (1979). Acoustic qualities of phonation in young hearing-impaired children. *Journal of Speech, Language, and Hearing Research*, *22*(2), 270-288.

Morton, J. &Jassem, W. (1965). Acoustic correlates of stress. *Language and Speech*, *8*, 159-179.

Mundy, I. R., & Carroll, J. M. (2012). Speech prosody and developmental dyslexia: reduced phonological awareness in the context of intact phonological representations. *Journal of Cognitive Psychology*, *24*(5), 560-581.

Nakatani, L. H., & Schaffer, J. A. (1978). Hearing’’words’’without words: Prosodic cues for word perception. *The Journal of the Acoustical Society of America*, *63*(1), 234-245.

Nakatani, L. H., O’Connor, K. D., & Aston, C. H. (1981). Prosodic aspects of American English speech rhythm. *Phonetica*, *38*(1-3), 84-105.

Nickerson, R. S. (1975). Characteristics of the speech of deaf persons. *Volta Review*, *77*(6), 342-62.

Nilsson, M., Soli, S. D., & Sullivan, J. A. (1994). Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise. *The Journal of the Acoustical Society of America*, *95*(2), 1085-1099.

Nowicki, S., & Duke, M. P. (1994). Individual differences in the nonverbal communication of affect: The Diagnostic Analysis of Nonverbal Accuracy Scale. *Journal of Nonverbal behavior*, *18*(1), 9-35.

Ohala, J. J. (1977). The physiology of stress. *Studies in stress and accent*, 145-168.

Owens, R. E. (2016). Language Development: An Introduction 9^{th} edition.

Parkhurst, B. G., & Levitt, H. (1978). The effect of selected prosodic errors on the intelligibility of deaf speech. *Journal of Communication Disorders*, *11*(2-3), 249-256.

Patil, H. A., Patel, T. B., Shah, N. J., Sailor, H. B., Krishnan, R., Kasthuri, G. R., ... & Kishore, S. P. (2013, November). A syllable-based framework for unit selection synthesis in 13 Indian languages. In *2013 International Conference Oriental COCOSDA held jointly with 2013 Conference on Asian Spoken Language Research and Evaluation (O-COCOSDA/CASLRE)* (pp. 1-8). IEEE.

Pearson. (2009). Advanced clinical solutions for WAIS–IV and WMS–IV. San Antonio, TX: Author.

Peppé, S., & McCann, J. (2003). Assessing intonation and prosody in children with atypical language development: the PEPS-C test and the revised version. *Clinical Linguistics & Phonetics*, *17*(4-5), 345-354.

Pisoni, D. B., & Sawusch, J. R. (1975). Some stages of processing in speech perception. In *Structure and process in speech perception* (pp. 16-35). Springer, Berlin, Heidelberg.

Priyanka,T (2018). Rhythm pattern of English in adult native Kannada speakers. *Unpublished Masters Dissertation, University of Mysore.*

Rajupratap, S. (1991) Production of word stress in children-3-4 years. *Unpublished Masters Dissertation, University of Mysore.*

Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition*, *73*(3), 265-292.

Reilly, A. P. (1979). *Syllable nucleus duration in the speech of hearing and deaf children*. University Microfilms International.

Rigault, A. (1962). Proceedings of the 4th International Congress of Phonetic Sciences, Helsinki, pp.735-748.The Hague: Mouton & Co.

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

Rohini, H (2006). Perception of stress in patients with Cerebro-Vascular Accident. *Unpublished Masters Dissertation, University of Mysore.*

Ross, E. D., Thompson, R. D., & Yenkosky, J. (1997). Lateralization of affective prosody in brain and the callosal integration of hemispheric language functions. *Brain and language*, *56*(1), 27-54.

Ruchi, A., Ghosh, K., & Savithri, S. R. (2007). Acoustic correlates of stress in Hindi. *In Proceedings of international symposium: Frontiers of research on speech and music* (pp. 281-283).

Russell, K. (1997). Intonation contours in English. Retrieved October 16, 2006, from http://www.umanitoba.ca/faculties/arts/linguistics/russell/138/sec3/inton.htm

Sallat, S., & Jentschke, S. (2015). Music perception influences language acquisition: melodic and rhythmic-melodic perception in children with specific language impairment. *Behavioural neurology, 2015*.

Sandner, G. W. (1981). Communication with a three-month old baby. In *Proceedings of the Thirteenth Annual Child Language Research Forum*.

Santhosh and Sahana (2012). Investigation of speech rhythm in individuals with stuttering: Understanding its nature, effect of fluency inducing conditions and the effect of treatment. *AIISH Research Fund Project.*

Savithri SR. (1987). Some acoustic and perceptual correlates of stress in Kannada. *In Proceedings of the National Symposium on Acoustics, Special edition* (Vol. 209), 42-45.

Savithri, Jayaram, Kedarnath and Sanjay (2006). Speech rhythm in Indo-Aryan and Dravidian languages. *AIISH Research Fund Project.*

Savithri, Johnsirani and Ruchi (2008). Speech rhythm in hearing impaired children. *AIISH Research Fund Project.*

Savithri, Maharani, Sanjay and Deepa (2007). Speech rhythm in Indian languages. *AIISH Research Fund Project.*

Savithri, S R., Sreedevi, N., Jayakumar, T., & Kavya, V. (2010). Development of speech rhythm in Kannada speaking children. JAIISH, 29 (1), 175-180.

Savithri, S. R. (1999a). Perception of word stress. *In Proceedings of the Madras India Regional Conference of the Acoustical Society of America (2)*, 110-113.

Savithri, Sreedevi and Kavya (2009a). Speech rhythm in Kannada speaking children. *In Proceedings of National Symposium on acoustics*, 1-6.

Savithri, Sreedevi and Kavya (2009b). Effect of gender on speech rhythm in Kannada speaking children. *In Proceedings of National Women's Science Congress*, 283-286.

Schlanger, B. B., Schlanger, P., & Gerstman, L. J. (1976). The perception of emotionally toned sentences by right hemisphere-damaged and aphasic subjects. *Brain and Language*, *3*(3), 396-403.

Shriberg, L. D., Kwiatkowski, J., Rasmussen, C., Lof, G. L., & Miller, J. F. (1992). *The Prosody-Voice Screening Profile (PVSP): Psychometric data and reference information for children* (No. 1, pp. 207-223). Tech. Rep.

Starkweather, C. W. (1987). *Fluency and stuttering*. Prentice-Hall, Inc.

Steinbeis, N., & Koelsch, S. (2008). Comparing the processing of music and language meaning using EEG and fMRI provides evidence for similar and distinct neural representations. *PloS one*, *3*(5), e2226.

Streeter, L. A. (1978). Acoustic determinants of phrase boundary perception. *The Journal of the Acoustical Society of America*, *64*(6), 1582-1592.

Stromberg, H., & Levitt, H. (1979). Multiple linear regression analysis of errors in deaf speech. *The Journal of the Acoustical Society of America*, *65*(S1), S69-S70.

Sussman, H., & Hernandez, M. (1979). A spectrographic analysis of the suprasegmental aspects of the speech of hearing-impaired adolescents. *Audiology and Hearing Education*, *5*, 12-16.

Thorsen, N. (1978). An Acoustical Investigation of Danish Intonation. *Journal of Phonetics*, *6*(3), 151-75.

Wade-Woolley, L. (2016). Prosodic and phonemic awareness in children's reading of long and short words. *Reading and Writing*, *29*(3), 371-382.

Wade-Woolley, L., & Heggie, L. (2015). Implicit knowledge of word stress and derivational morphology guides skilled readers' decoding of multisyllabic words. *Scientific Studies of Reading*, *19*(1), 21-30.

Wells, B., & Peppé, S. (2003). Intonation abilities of children with speech and language impairments. *Journal of Speech, Language, and Hearing Research*, 46 (1), 5-20.

Westin, K., Buddenhagen, R. G., & Obrecht, D. H. (1966). An experimental analysis of the relative importance of pitch, quantity, and intensity as cues to phonemic distinctions in southern Swedish. *Language and speech*, *9*(2), 114-126.

Whalley, K., & Hansen, J. (2006). The role of prosodic sensitivity in children's reading development. *Journal of Research in Reading*, *29*(3), 288-303.

Williams, B. (1985). Pitch and duration in Welsh stress perception: The implications for intonation. *Journal of Phonetics*.

Wilson, R. H. (2003). Development of a speech-in-multitalker-babble paradigm to assess word-recognition performance. *Journal of the American Academy of Audiology*, *14*(9), 453-470.

Wingate, M. E. (1984). Stutter events and linguistic stress. *Journal of Fluency Disorders*, *9*(4), 295-300.

Wood, C., & Terrell, C. (1998). Poor readers' ability to detect speech rhythm and perceive rapid speech. *British Journal of Developmental Psychology*, *16*(3), 397-413.

APPENDICES

APPENDIX - I

The developed test for prosody skills was administered on native Kannada speaking typically developing children in the age range of 5-8 years. This test is intended for use among the clinical population. Test instructions and procedure to administer the test are as follows:

Task 1 - Stress task

Instructions- *"Let us play a game. I will be placing this headphone on your ears and you will hear two /ma/ sounds. Listen to the /ma/ sounds carefully and indicate if both are same or different".*

Adequate number of trials can be provided until the participant understands the task.

If the participant is unable to follow the task, the following procedure can be followed:

1. ***Instructions*** - *"Let us play a game. I am going to produce two /ma/ sounds now and you have to tell me whether both the sounds are perceived as same or different."*
2. Clinician should produce two /ma/ syllables with a pause of approximately 500-1000ms in between.
3. The clinician should instruct the participant to produce the perceived sounds in similar manner. This step is carried out to ensure that the participant has understood the task and stimuli.
4. Participants are permitted to produce the /ma/ syllables with a change in any combination - frequency/ intensity/ duration (We know that it is difficult to differentiate specific changes in different acoustic parameters and a single

parameter cannot be focused during natural speech production of a stressed syllable).

5. The clinician needs to remember the focus here is to make the participants indicate that there is difference / no difference between the two /ma/ syllables presented.
6. Once the participant has understood the task and method of response, experimental testing should be carried out with the stimuli tokens.

Similar instructions and procedure should be followed in the non-verbal stress task also. The stimuli for non-verbal stress task consist of piano tones instead of syllable /ma/.

Task 2 - Intonation task

Instructions - "*Let us play a different game. I will be placing this headphone on your ears and you will hear different tones of /a/. Listen to the stimuli carefully and imitate the perceived /a/ in a similar manner*"

If the participant is unable to follow the task, following procedure can be carried out:

7. The clinician can produce different intonation patterns with vowel /a/ and ask the child to imitate the same along with the clinician.
8. Once the participant is able to produce the target intonation patterns along with the clinician, the clinician can fade off his/ her participation during production.
9. Now clinician should present the stimuli tokens and instruct the participant to produce the perceived intonation patterns.

The same instructions and procedure should be followed for non-verbal intonation task also. The stimuli for non-verbal stress task consist of tone glides.

Task 3 - Rhythm task

Verbal rhythm task

Instructions - "*I will be placing headphone on your ears like before and you will hear a string of sound/ma/. I will indicate once the sound has ended by tapping on your hand. Listen carefully to the sound /ma/ and imitate whatever you heard in a similar manner.*"

If the participant is not able to understand the task, a similar procedure (refer to Task 2) like that of intonation task can be followed. The clinician and the participant can imitate the perceived tokens together initially. After the participant is able to imitate on his/ her own, testing can be carried out.

Non-verbal rhythm task

Instructions - *"I will be placing this headphone on your ears like before and you will hear sounds like tak….tak…tak….(onomatopoeic production). As soon as you listen to the sound you have to tap on this place (point to the response pad of the laptop/ PC; the stimuli are presented using the computerized application). After some time, you will not be hearing sounds but still continue tapping in the same pace as earlier. You can stop the task after I indicate it to you. Now, tap exactly at the same time you hear the sound and let us see if you can tap without missing each sound"*

The testing duration of each stimuli condition (20 BPM paced, 20 BPM unpaced, 40 BPM paced and 40 BPM unpaced conditions) is 1-minute. The clinician should indicate the participant to stop tapping in unpaced condition after 1-minute of testing is over (as the participant will not be aware if he/ she has to continue tapping since there will not be any stimuli in the unpaced condition).

If the participant is unable to follow and perform the task, the following procedure can be carried out:

- The clinician and the participant can tap the response pad (of the laptop) simultaneously in accordance with stimuli.
- The computerized task should be stopped once the participant has understood the task.
- The clinician should re-start the computerized task for testing.

Steps to perform the computerized task are as follows:

- Open the audio files of metronome beats (20 BPM should be tested first followed by 40 BPM audio file) through any music player (widows media player, VLC player, etc,.) available in a PC / Laptop.
- Open the computerized task and select the "*Start*" button in it.
- Now instruct the participant to tap the response pad of the laptop (click the mouse in case of PC) along with the perceived beats, after placing the cursor on "*Click Me*" icon in the computerized task.
- After 1 minute of testing with the stimuli, close the audio file and ask the participant to continue tapping the response pad for 1 more minute. The former will be the paced (with stimuli) testing and latter will be the unpaced (without stimuli) testing.
- After 2-minute testing for the first BPM condition (that is, 20 BPM condition), click the "*End*" icon on the task display.
- Once the "*End*" icon is selected, the time of tap/ click (in laptop/ in a PC) according to each tap/ click will be displayed as an output in the task.
- Select the option "*Download*" on the display and save the output in a folder to document the time of tap for each token.
- Similar procedure should be followed for the 40 BPM condition.

APPENDIX – II

Inter-judge reliability Kappa values obtained in intonation task

Group I - Typically developing children

Age range: 5-6 years

Stimuli token	Kappa value
HVT1	1.00
HVT2	0.87
HVT3	0.65
FVT1	0.92
FVT2	0.80
FVT3	0.92
RVT1	0.83
RVT2	1.00
RVT3	1.00
LVT1	1.00
LVT2	1.00
LVT3	1.00
HNVT1	0.85
HNVT2	0.90
HNVT3	0.46
FNVT1	0.79
FNVT2	0.93
FNVT3	0.78
RNVT1	0.83
RNVT2	0.51
RNVT3	0.78
LNVT1	1.00
LNVT2	1.00
LNVT3	1.00

(Note: T1, T2, T3 - indicate the token numbers; V & NV - indicates verbal and non-verbal stimuli; H, F, R & L - indicates high steady, falling, rising & low steady intonation patterns respectively)

Age range: 6-7 years

Stimuli token	Kappa value
HVT1	0.75
HVT2	0.83
HVT3	1.00
FVT1	0.72
FVT2	0.81
FVT3	0.51
RVT1	0.61
RVT2	1.00
RVT3	0.48
LVT1	1.00
LVT2	1.00
LVT3	1.00
HNVT1	0.47
HNVT2	1.00
HNVT3	1.00
FNVT1	0.84
FNVT2	1.00
FNVT3	1.00
RNVT1	1.00
RNVT2	1.00
RNVT3	1.00
LNVT1	1.00
LNVT2	1.00
LNVT3	1.00

Age range: 7-8 years

Stimuli token	Kappa value
HVT1	0.63
HVT2	1.00
HVT3	1.00
FVT1	0.66
FVT2	0.60
FVT3	0.65
RVT1	1.00
RVT2	1.00
RVT3	0.78
LVT1	1.00
LVT2	1.00
LVT3	1.00
HNVT1	1.00
HNVT2	1.00
HNVT3	1.00
FNVT1	0.91
FNVT2	0.65
FNVT3	1.00
RNVT1	1.00
RNVT2	1.00
RNVT3	1.00
LNVT1	1.00
LNVT2	1.00
LNVT3	1.00

Group II - Children with SLI

5-8 years

Stimuli token	Kappa value
HVT1	1.00
HVT2	1.00
HVT3	1.00
FVT1	1.00
FVT2	0.91
FVT3	1.00
RVT1	1.00
RVT2	1.00
RVT3	1.00
LVT1	1.00
LVT2	1.00
LVT3	1.00
HNVT1	1.00
HNVT2	1.00
HNVT3	1.00
FNVT1	1.00
FNVT2	1.00
FNVT3	1.00
RNVT1	1.00
RNVT2	1.00
RNVT3	1.00
LNVT1	1.00
LNVT2	1.00
LNVT3	1.00

APPENDIX – III

Inter-judge reliability Kappa values obtained in verbal rhythm task

Group I - Typically developing children

Age range: 5-6 years

Stimuli token	Kappa value
F1T1	1.00
F1T2	1.00
F1T3	1.00
F2T1	0.5
F2T2	0.5
F2T3	1.00
F3T1	0.88
F3T2	1.00
F3T3	0.70
F4T1	0.81
F4T2	0.73
F4T3	0.81
F5T1	0.87
F5T2	1.00
F5T3	0.69
I1T1	1.00
I1T2	1.00
I1T3	1.00
I2T1	1.00
I2T2	1.00
I2T3	1.00
I3T1	0.82
I3T2	1.00
I3T3	1.00
I4T1	0.6
I4T2	1.00
I4T3	0.81
I5T1	1.00
I5T2	1.00
I5T3	1.00

Stimuli tokens	Kappa value
FI1T1	1.00
FI1T2	1.00
FI1T3	1.00
FI2T1	1.00
FI2T2	1.00
FI2T3	1.00
FI3T1	1.00
FI3T2	0.64
FI3T3	0.86
FI4T1	1.00
FI4T2	1.00
FI4T3	1.00
FI5T1	1.00
FI5T2	1.00
FI5T3	1.00

(Note: T1, T2 & T3 - token numbers; F1, F2,…F5 - frequency only manipulation condition; I1, I2…I5- intensity only manipulation condition; FI1, FI2,…FI5 - combined frequency-intensity manipulation condition)

Age range: 6-7 years

Stimuli token	Kappa value
F1T1	1.00
F1T2	1.00
F1T3	1.00
F2T1	1.00
F2T2	1.00
F2T3	1.00
F3T1	1.00
F3T2	1.00
F3T3	1.00
F4T1	1.00
F4T2	1.00
F4T3	1.00
F5T1	0.5

F5T2	0.76
F5T3	1.00
I1T1	1.00
I1T2	1.00
I1T3	1.00
I2T1	1.00
I2T2	1.00
I2T3	1.00
I3T1	1.00
I3T2	1.00
I3T3	0.69
I4T1	0.5
I4T2	0.5
I4T3	1.00
I5T1	0.76
I5T2	1.00
I5T3	1.00
FI1T1	1.00
FI1T2	1.00
FI1T3	1.00
FI2T1	1.00
FI2T2	1.00
FI2T3	1.00
FI3T1	1.00
FI3T2	0.46
FI3T3	1.00
FI4T1	1.00
FI4T2	1.00
FI4T3	1.00
FI5T1	1.00
FI5T2	0.5
FI5T3	0.5

Age range: 7-8 years

Stimuli token	Kappa value
F1T1	1.00
F1T2	1.00
F1T3	1.00
F2T1	1.00
F2T2	1.00
F2T3	1.00
F3T1	1.00
F3T2	1.00
F3T3	1.00
F4T1	1.00
F4T2	1.00
F4T3	0.6
F5T1	1.00
F5T2	0.65
F5T3	1.00
I1T1	1.00
I1T2	1.00
I1T3	1.00
I2T1	1.00
I2T2	1.00
I2T3	1.00
I3T1	1.00
I3T2	1.00
I3T3	1.00
I4T1	1.00
I4T2	0.5
I4T3	0.61
I5T1	0.52
I5T2	1.00
I5T3	1.00
FI1T1	1.00
FI1T2	1.00
FI1T3	1.00
FI2T1	1.00
FI2T2	1.00
FI2T3	1.00
FI3T1	1.00

FI3T2	1.00
FI3T3	1.00
FI4T1	1.00
FI4T2	1.00
FI4T3	0.5
FI5T1	1.00
FI5T2	0.8
FI5T3	1.00

Group II - Children with SLI

Age range: 5-8 years

Stimuli token	Kappa value
F1T1	1.00
F1T2	1.00
F1T3	1.00
F2T1	1.00
F2T2	1.00
F2T3	1.00
F3T1	0.5
F3T2	1.00
F3T3	1.00
F4T1	1.00
F4T2	1.00
F4T3	1.00
F5T1	1.00
F5T2	1.00
F5T3	1.00
I1T1	1.00
I1T2	1.00
I1T3	1.00
I2T1	1.00
I2T2	0.83
I2T3	1.00
I3T1	1.00
I3T2	1.00
I3T3	0.46
I4T1	1.00

I4T2	1.00
I4T3	1.00
I5T1	1.00
I5T2	1.00
I5T3	1.00
FI1T1	1.00
FI1T2	1.00
FI1T3	1.00
FI2T1	1.00
FI2T2	1.00
FI2T3	1.00
FI3T1	1.00
FI3T2	0.76
FI3T3	1.00
FI4T1	1.00
FI4T2	1.00
FI4T3	1.00
FI5T1	1.00
FI5T2	1.00
FI5T3	1.00

APPENDIX - IV

The audio files of developed stimuli tokens to assess prosody are provided in the folder named “Stimuli” in the same CD of the project report.

APPENDIX – V

Mean/ median scores of Group I (TDC) and Group II (CWSLI) participants in stress, intonation and rhythm tasks are provided as follows:

1. Stress task (mean scores)

Task	**Group I (TDC)**			**Group II (CwSLI)**
	5-6 years	6-7 years	7-8 years	5-8 years
Verbal	30.90	47.44	73.47	46.10
Non-Verbal	66.01	72.64	74.54	62.75

2. Intonation task (median scores)

Task	**Intonation patterns**	**Group I (TDC)**			**Group II (CwSLI)**
		5-6 years	6-7 years	7-8 years	5-8 years
Verbal	High steady	3	3	3	3
	Falling	2	2	2.5	0
	Rising	3	3	3	0
	Low steady	3	3	3	3
Non-Verbal	High steady	3	3	3	3
	Falling	2	3	3	1.5
	Rising	3	3	3	0
	Low steady	3	3	3	3

3. Rhythm task

Verbal rhythm task (mean scores)

Stimuli manipulation	Group I (TDC)			Group II (CwSLI)
	5-6 years	6-7 years	7-8 years	5-8 years
Frequency	18.23	23.90	24.93	15.30
Intensity	18.93	20.39	22.33	13.40
Frequency-Intensity	20.46	24.00	25.63	19.50

Non-Verbal rhythm task (mean scores)

Stimuli Condition	Group I (TDC)			Group II (CwSLI)
	5-6 years	6-7 years	7-8 years	5-8 years
P-20	3.65	2.46	3.05	4.21
P-40	1.89	1.99	1.58	1.96
UP-20	4.22	2.84	3.38	4.92
UP-40	2.27	2.14	1.87	2.54

(**Note:** P-20= Paced 20 BPM; P-40= Paced 40 BPM; UP-20= Unpaced 20 BPM; UP-40= Unpaced 40 BPM)