PROJECT REPORT

Perception and Production of Prosody in Children with Hearing Impairment

---

AIISH Research Fund Project No: SH/CDN/ARF-SLP-KY/2018-19

Investigators:

Dr. Yeshoda K.

Dr. Sreeraj Konadath

Sanction No. SH/CDN/ARF-SLP-KY/2018-19

Personnel	Ms. Revathi R Research Officer
Total Fund	Rs. 4,93,000/-
Project duration	12 months
Principal Investigator	Dr. Yeshoda K Associate Professor in Speech Sciences, Department of Speech-Language Sciences All India Institute of Speech and Hearing Mysuru– 570006, Karnataka, India
Principal Investigator	Dr. Sreeraj Konadath Asst. Professor in Audiology, Department of Audiology All India Institute of Speech and Hearing Mysuru– 570006, Karnataka, India

Acknowledgments

The investigators would like to acknowledge the Director, All India Institute of Speech and Hearing, Mysuru, for being pivotal in granting the funds for the study. Our heartfelt gratitude is also extended to all the participants for their kind cooperation.

Contents

List of Figures

List of Tables

ABSTRACT

The project aimed to study the production and perception of prosody in children with hearing impairment and compare with typically developing children matched for their language ages. The present study was carried out in two phases. The first phase was pertaining to the training on perception and production of prosody, and the second phase was pertaining to the testing of the perception and production of prosody in terms of stress, intonation, and rhythm. A set of 24 Malayalam sentences was formulated that matches the language age of 3years to 7years, and picture representations of these sentences were made. The training phase included the participants listening to 14 sentence stimuli and responding by selecting appropriate pictures depicting particular prosody. All the participants received a prosody training of 180 minutes and then completed the testing. The testing phase was done using the 10 test stimuli. The participants were instructed to give the appropriate response after listening to the stimuli for the perception task. In the production task, they were instructed to repeat the prosody-loaded sentences presented in a randomized order. The responses were audio-recorded using a digital recorder and were then subjected to acoustic analysis. The measurement and comparison of the parameters of stress, intonation, and rhythm of children with hearing impairment and their language age-matched typically developing children was carried out. There were significant differences found between the groups on perception and production of prosody. A chronology of the development of rhythm was derived from the results of the study. The study results also indicated that prosody training will help increase the clarity and intelligibility and thereby naturalness of the spoken language skill of children with hearing impairment. The current study ascertains the need for conventional speech-language and auditory training to expand and give equal importance to segmental and supra-segmental training in children with hearing impairment.

CHAPTER 1
INTRODUCTION

Communication is made successful with the help of segmental as well as supra segmental features of speech. The segmental features comprise the constituent phonemes and their distinctive features (vowels and consonants), while the supra segmental features include the variations brought about in the fundamental frequency, intensity, and duration. The suprasegmental features improve speech understanding. They enable the receiver to have a perfect interpretation of the sender's stimuli, helping to differentiate between a statement and a question, emotions, and even linguistic units like a noun and a verb. Both types of information are simultaneously encoded in the produced speech and are necessary for effective communication (Abercrombie, 1967; Pisoni, 1997). Over the years, there has been a shift of focus from the purely segmental aspects of speech to the communication process as a whole, in which the supra segmental parameters play a significant role.

Prosody can be described as the melody and rhythm of speech. Operationally, prosody can be defined as the suprasegmental features of speech conveyed by the parameters of the fundamental frequency, intensity, and duration; such suprasegmental features include stress, intonation, tone, and tone duration (Kent & Kim, 2008). The prosodic structure development in children is essential as it carries different functions like linguistic and affective; thus becomes a crucial factor in the transfer of meaning of information and, subsequently, the intelligibility. Typically developing, normal-hearing children start picking up the prosody in the motherese during the initial stages of language development (Fernald, 1985).

For the language acquisition in typically developing children, both segmental and supra-segmental features of the speech play a significant role. According to the study done by Werker and Yeung (2005), infants start picking up the acoustic cues of stress, intonation, and rhythm, to understand the meaning of the utterance, much before starting to grasp and categorize the phonemes of their particular language. Children learn the phrase boundaries and the foot structures of their mother tongue, perceiving the prosodic patterns of the mother and the caregivers (Thiessen & Saffran, 2003; Hirsch-Pasek et al., 1987). Infants attend to the exaggerated, child-directed intonation patterns in their immediate environment, enabling them to identify word boundaries and then acquire their initial words (Seidl & Cristi`a, 2008; Seidl & Johnson, 2008; Swingley, 2009). Multiple language exposure has also resulted in children coming up with cross-linguistic prosody patterns in their early productions. Typically developing children also exhibit cross-linguistic prosodic patterns in their speech productions.

The attention a child gives to the perception of prosodic patterns of the speakers forms the fundamental steps of language acquisition (Thiessen & Safran, 2003; Thiessen et al., 2005; Werker & Yeung, 2005; Massicotte-Laforge & Shi, 2015). The perception of prosodic patterns will be disrupted in children with hearing impairment. This deficit in perception of prosody gets manifested as monotonous speech production as well as delay in language acquisition in children with hearing loss (Lederberg et al., 2000). For long, studies have been conducted on suprasegmental aspects of language to examine the efficacy of the amplified speech in conveying the prosodic cues that help in better-spoken word recognition. Studies were conducted by Erber and colleagues (1979) in an effort to establish a relationship between the pure tone thresholds and the perception of suprasegmental (syllable stress) and segmental characters in children. Results revealed that perception of suprasegmental features was better than segmental features' perception (Erber 1972a,b; 1979; Hack & Erber, 1982).

Similar studies were done by Boothroyd (1994), assessing the perception of segmental and suprasegmental characteristics in children with hearing impairment. The subjects were sub-grouped according to their levels of hearing. The results revealed that the overall perception of suprasegmental tasks was better than the performance on segmental tasks. Recently, many studies were done on children with a hearing loss of severe to profound degree, which uses different types of amplification devices like hearing aids (HA) or cochlear implants (CI).

From the studies, it was inferred that differences in the type of amplification could cause perceptual differences of suprasegmental cues (Kong et al., 2005). Many studies reported that the perception of segmental aspects (Eg: word recognition, sentence recognition) was better with cochlear implant, while with hearing aids, a better perception of the suprasegmentals (Eg: pitch, stress, emotions) was observed (Blamey et al., 2001; Most & Peled, 2007; Most et al., 2011). Hearing aids can transfer the suprasegmental cues efficiently, but the amount of spectral content a hearing aid can convey is majorly limited to the hearing thresholds of the user. On the contrary, the cochlear implants can do only minimal processing of low-frequency information, and thus leaving out the acoustic information (pitch, intensity, duration) that are significant for the transfer of suprasegmental information (Carroll & Zeng, 2007; Most & Peled, 2007; Nittrouer & Chapman, 2009; Most et al., 2011). At the same time, cochlear implants can convey the temporal envelope information efficiently such that the vowel and consonant cues are transferred with utmost precision (Van Tasell et al., 1992; Shannon et al., 1995).

A review of studies related to the perception and production of stress, rhythm, and intonation in children with hearing impairment is discussed below.

1.1 Studies on Speech Stress

In many languages like English, during verbal communication, the syllables in polysyllabic words are given different levels of prominence and are perceived differently with each combination of prominence. This supra segmental feature called stress emphasizes the components of speech: a whole word, a part of the word, or a syllable of the word and plays a significant role in the understanding of the speech (Raphael, Borden, & Harris, 2007). Over the years, authors came up with psychophysiological and physiological viewpoints of stress. Stress was defined as a greater speaking effort (Lehiste, 1970); increased sub-glottal pressure (Fonagy, 1966); louder expressions (Bloomfield, 1933; Trager & Smith, 1951);

prominence imposed within utterances (Bolinger, 1958). According to the degree of emphasis, the stress in speech is categorized as primary, secondary, and unstressed (Hayes, 1995; Echols & Newport, 1992; Arciuli, Monaghan, & Seva, 2010). The lexical and emphatic are the two major types of stress explained in the literature. Lexical stress is the emphasis given on a syllable in a word that can change the meaning of the word (Pisoni & Remez, 2005; Clerck, Verhoeven, Gillis, Pettinato, & Gillis, 2019). Emphatic stress conveys different ideas according to the position of the stressed word in the phrase or a sentence (Asher, 2013).

1.1.1 Acoustic cues for stress

The acoustic realizations of stress in the time energy envelope among different languages were studied by researchers around the globe. Three basic physical correlates of stress: intensity, duration, and fundamental frequency were identified to have varying prominences in stress expression in different languages (American English, Swedish, French, Polish, Estonian, Serbo-Croatian, Indian languages, etc.). Lieberman (1960) found stress in American English to be evident as higher fundamental frequency, peak amplitude, and longer duration. Rigault (1962) observed that variations in frequency and duration marked stress in the French language. In Swedish, the fundamental frequency was found to be the significant marker of stress (Westin, Buddenhagen, & Obrecht, 1966). In the Polish language - fundamental frequency (Jassem, 1959) and duration; Estonian language – duration alone (Lehiste, 1968a); Serbo-Croatian language – duration (Lehiste & Ivic, 1973) were observed to produce stress.

Similar findings were reported by studies carried out in Indian languages, both Indo-Aryan and Dravidian languages suggesting that the acoustic realizations of stress vary across languages. Emphatic stress was observed to be the primary type of stress in Indian languages, especially the Dravidian languages (Mohanan, 1989). The major acoustic cue for stress was perceived as intensity in Telugu (Srinivas, 1992); both duration and intensity in Tulu (Narra, Teja, Sneha & Dattatreya, 2012); both duration and fundamental frequency in

Hindi (Ruchi, Ghosh, & Savitri, 2007; Dhillon, 2010). The duration was proposed to be carrying the major acoustic cue for emphatic stress in Kannada (Rajupratap, 1991; Savithri, 1999), Konkani (Kumar & Bhat, 2009), and Tamil (Keane, 2006). Duration was the most significant contributor to the emphatic stress in the Malayalam language by Irfana, Rofina & Sreedevi (2014).

Even at 1-4 months, infants are observed to have the ability to discriminate syllables only in stress placement (Spring & Dale, 1977). The infants were observed to identify novel words based on their stress patterns and shifts in the position of the stress in the same word. This suggests that even at an early age, infants perceive and learn novel words along with their prosodic attributes. Infants as young as 17-months old use the stress pattern along with the phoneme sequences during their early word acquisition stages (Estes & Bowen, 2013). Thus, the perception of prosodic cues can be regarded as a significant element in early language acquisition. With prosodic cues, children learn to assign meanings to novel words, attaching word boundaries in the perceived speech (Newman, Ratner, Jusczyk, Jusczyk, & Dow, 2006).

Children's use of stress in their speech also starts at a young age. Brown (1973) reported that children as young as 18 months used contrastive stress in their speech while introducing a new concept. Evidence suggests that children as young as 2-3 years of age produce stress in their speech, even with correct stress placement (Klein, 1984; Macwhinney & Bates, 1978). In Indian languages, the production of stress by Kannada speaking typically developing children in the age range of 3-4 years was studied by Raju Pratap (1991). Results indicated that the production of word stress was marked by greater loudness and increased duration, and the production increased from 3years to 4 years. The female children in the age range of 3.10 – 4 years were able to obtain a 100% score on the production of stress. The author attributed the increase in stress production to physiological development and an increase in vital capacity. Jaya (1992) found a developmental trend in the production of stress in Tamil-speaking, typically developing children in the age range of

2-8years of age. The author also reported that the capacity to perceive, imitate, and produce stress use increases as a function of age.

However, there is no clear normative found on the development of stress throughout childhood. Also, significantly less is known about the contribution of the different parameters of stress (intensity, frequency, and duration) at different age levels (Ballard, Djaja, Arciuli, James, & Doorn, 2012). As infants grow older, they start relying on the other rules of language and give less importance to stress cues (Vavatzanidis, Mürbe, Friederici, & Hahne, 2016). Nonetheless, stress perception and production are considered the fundamental step critical for early and later language development (Aslin & Newport, 2009).

1.1.2 Perception and production of stress in individuals with hearing impairment

The importance of prosodic cues in language acquisition becomes evident in children who are less sensitive to such cues. In children with congenital severe to profound hearing impairment, only limited access to the prosodic cues in the speech signal is available (Clement, 2004). The ability to detect the spectral cues in a speech stream is minimal in severe to profound hearing loss, and thereby the major acoustic cues they use are the gross variations in the acoustic pattern (Erber, 1979). While for the perception of syllable stress in a word or word stress in a sentence, the signal's time-energy envelope and/or the fundamental frequency (with respect to the language spoken) is significant (Cohen-Licht & Most, 2000). Despite this, researchers have pointed out that majority of the children with hearing impairment have a residual hearing at the low frequency, and thereby the suprasegmental cues like duration, intensity and to an extent, the frequency variations should be perceivable to them (Boothroyd, 1984; Most, 1999). Studies suggest that while perceiving supra segmental cues, persons with hearing loss find stress cues from the time-energy envelope more accessible than the intonation cues from fundamental frequency (Raphael et al., 2007).

Reviewing the studies on the perception of prosody through hearing aids, a mixed set of results showcasing positive and negative results can be found. Some studies reported that children with severe- profound loss fitted with hearing aids had difficulty perceiving stress patterns (Jackson & Kelly, 1986; Osberger & McGarr, 1982). Similar results were found by researchers who studied other suprasegmental features like intonation (Stark & Levitt, 1974). Studies reported that subjects with severe to profound hearing loss could perceive the stress patterns only when there was an amplitude variation between the stressed and unstressed syllables (Rubin-Spitz & McGarr, 1986; Most, 2015). On the contrary, studies also reported that hearing aid users had a near-normal perception of prosodic characters like syllable stress and intonation patterns (Most, 1999). Similar results were reported by Boothroyd and Eran (1994), Most and Aviner (2009), and Laugen, Jacobsen, Rieffe, and Wichstrøm (2016), where the hearing-impaired group was able to perceive the prosodic cues. Engen, Engen, Clarkson, & Blackwell (1983) reported that hearing aid users were able to detect the time–energy envelope and discriminate the variations in the envelope.

The limited exposure to stress cues because of the auditory deficit limits the expression of the stress in children with hearing impairment. Studies report that the demarcation between stressed and unstressed syllables is very minimal in the speech productions of children with hearing impairment (Angelocci, 1962; John & Howarth, 1965; Nickerson et al., 1974). Boone (1966) reported that deaf speakers produce only stressed syllables, while Boothroyd et al. (1974) reported individuals with hearing impairment to produce unstressed syllables twice longer than the average hearing individuals. Graddol (1991) reported that individuals with hearing impairment produce too many stressed syllables in their speech. Studies reported that individuals with hearing impairment could discriminate between stressed and unstressed words only when they differed on amplitude (Rubin-Spitz & McGarr, 1986).

In a study done by Clement, Os, and Beinum (1996) comparing the babbling expressions of 5-10 months old typically developing infants and infants with profound hearing impairment, it was observed that both the groups of infants had differences at durational and syllabic levels, which was concluded as a consequence of lack of auditory feedback. Most (1999) found that children with hearing impairment increased the F0 and amplitude of stressed words in comparison with that of unstressed words. O'Halpin (1993, 2001) found that after some training, one child with hearing impairment could produce stress in his speech by producing appropriate lengthening of the target syllables. The studies done on the production of stress in children with hearing impairment are scarce, especially in Indian languages.

1.2 Studies on Speech rhythm

Speech rhythm basically can be described as the recurrence of a particular pattern of some parameter in an otherwise constant temporal environment (Gibbon & Gutt, 2001). Cumming (2011) defined speech rhythm as the consistency that is comprehensible and created by the grouped prominences (including different speech acoustic signal characteristics) and affected by the perceiver's dialect. Rhythm can be defined as an isochronous recurrence of some type of speech unit (Pike, 1946; Abercrombie, 1964). This patterning/recurrence can be assigned to value sequences in a high-low pitch, long-short syllable, or even vowel-consonant segmentation. Nakatani, O'Connor, and Aston (1981) found some factors that influence rhythm as syllable duration, word stress, word size, syllable stress, position within the word, syntactic content, and metrical feet. Wiget et al. (2010) spoke of more general factors like the rate of speech of the individual, the dialect used, environment, type of speech act, and type of stimuli likely to influence the speech rhythm. Duffy (2005) proposed physical factors like accurate articulatory movements and transitions, coordination between the subsystems, tone, and strength of articulatory muscles to influence the production of speech rhythm.

1.2.1 Acoustic Cues for Speech Rhythm

Though many acoustic parameters like pitch, intensity, duration and the transition of the second formant frequency account for the varied rhythm patterns in different languages, the role of timing has been studied extensively. In general, the perceptual distinction between the stressed and unstressed syllables characterizes each language with its contrastive rhythm. Contrastive rhythm of a particular language imparts an aesthetic appeal as well as naturalness to the speech (Amulya, 2013). According to the patterning of the speech units producing contrastive rhythm, the languages were divided into stress-timed, syllable-timed, and mora-timed languages (Abercrombie, 1965; Ladgefod, 1975). In stress-timed languages, there is a high contrast between the stressed and unstressed syllables like in the English language, while in syllable-timed languages, the durations of stressed and unstressed syllables are equal, like in the Spanish language. In mora-timed languages, morae are syllables that consist of one short vowel and any preceding onset consonants. Consecutive morae are said to be almost equal in duration. Thus, mora-timed languages are more similar to syllable-timed languages than stress-timed languages.

Rhythm Class Hypotheses (RCH) proposed by Pike (1946) and Abercrombie (1967) states that every language can be assigned to one of the three rhythm classes: stress-timed, syllable-timed, and mora-timed. A rhythm class is defined by assigning durations to selected speech units specific to the language. The parameter for rhythm classification was listed by several authors like syllable duration (Abercrombie, 1967) and inter-stress interval (Roach, 1982). Ramus, Nesper, and Mehler (1999) proposed %V (the proportion of time taken by all the vocalic segments in the sentence, omitting the boundaries), ∆V (the standard deviation of the vocalic intervals), and ∆C (the standard deviation of consonant intervals) as standardized parameters to categorize rhythm class. Grabe and Low (2000) devised normalized Pairwise Variability Index (nPVI) and raw Pairwise Variability Index (rPVI), where the duration of the vocalic and intervocalic intervals was used as the raw data. The indices expressed the level of variability in successive measurements, using the given equations (1) & (2):

$$nPVI = 100 \times \frac{\left(\sum_{k=1}^{m-1} \left| (d_k - d_{k+1})/((d_k + d_{k+1})/2) \right| \right)}{(m-1)} \qquad \text{Equation (1)}$$

$$rPVI = \left(\sum_{k=1}^{m-1} \left| d_k - d_{k+1} \right| \right) / (m-1) \qquad \text{Equation (2)}$$

Where,

m—number of intervals in an utterance for which PVI is calculated,

d—duration of the kth interval.

Based on the PVI values for vocalic (nPVI) and intervocalic (rPVI) segments, rhythm patterns were classified as shown in Table 1.1 (Low, Grabe & Nolan, 2000). The PVI above 50 was considered high, and a PVI below 50 was considered low (Merin, 2016).

Table 1.1 Rhythm pattern classification based on vocalic and intervocalic segments

	Stress –timed	**Syllable-timed**	**Mora-timed**
nPVI	High	High	Low
rPVI	High	Low	High

Note: nPVI- normalized Pairwise Variability Index (nPVI); rPVI- raw Pairwise Variability Index

The speech rhythm classification works on the phonological pattern of the language (Grabe & Low, 2002). In the Indian context, studies were carried out to study the cross-linguistic variations of rhythm in Indian languages (Savithri, Jayaram, Kedarnath, & Goswami, 2006; Savithri, Maharani, Goswami, & Deepa, 2007). The authors found that Kannada is mora-timed, with nPVI (vocalic) mean of 46.18 and rPVI (intervocalic) mean of 46.95 (Savithri et al., 2006). Similarly, Savithri et al. (2007) conducted a cross-language study. They found that Assamese, Punjabi, Kannada, and Telugu are mora-timed, whereas Hindi, Bengali, Malayalam, Tamil, and Kashmiri are syllable-timed languages. The developmental rhythm studies carried out on Kannada-speaking children revealed mixed results with different age groups having different rhythm classes, majorly syllable-timed, mora-timed, and unclassified (Savithri, Sreedevi, Deepa, & Aparna, 2011). A high value of nPVI and a low value of rPVIwas found in Kannada-speaking children in the age range of 5-10 years (Savithri, Johnsirani, & Ruchi, 2008).

In most of the clinical populations with a developmental communication disorder, the rhythm is affected. Rhythm characters setting the boundaries between vowel and consonant segments is crucial in the acquisition of language in the early developmental years. Mehler, Jusczyk, Lambertz, Halsted, Bertoncini, and Amiel-Tison (1988) reported that newborns depended on rhythm patterns to discriminate between their native and other languages.

1.2.2 Perception and Production of Speech Rhythm

Despite the apparent need for exploring the area of perception and production of rhythm, studies conducted in this area have been few and have yielded mixed results. The early studies conducted by Haal and Jastrow (1885); Bolton (1894) used psychophysical parameters to study the perception of speech rhythm. The stimuli used were clicks, and most of the subjects were not able to maintain the requested organization of the stimuli for a longer duration and went back to their preferred pattern of organization. Odekar (2001) studied the perception of rhythm in normal-hearing adults, using the talas of Carnatic music. Each participant was presented with rhythms monoaural and was instructed to perform table tapping in accordance with the rhythm heard. The participants were able to replicate the rhythm taps accurately in both monaural and dichotic conditions.

The pioneer studies pertaining to the perception of speech rhythm include the study done by Scott, Israd, and de Boysson-Bardies (1985), which revealed that the tapping pattern of the participant in response to French/English stimuli was not heavily influenced by their native language. While, at the same time, Miller (1984) found differences in the way the participants classified the languages into stress-timed and syllable-timed, depending on the native language of the participant. Nazzi, Bertoncini, and Mehler (1998a) conducted a study on French-speaking newborns with sets of different sentences in foreign languages. They found that the language discrimination in their participants was based on rhythm and rhythmic classes. Similarly, Ramus, Hauser, Miller, Morris, and Mehler (2000) revealed in their study that French newborns were able to discriminate between stress-timed and mora-timed languages, even when the other non-rhythmic speech cues which could have facilitated the discrimination between the languages were controlled. Studies comparing the speech rhythm of children and adults speaking the same language revealed that children prefer to speak in an equally timed pattern of rhythm, irrespective of language (Arvaniti, 2009). It indicates that equal timing patterns dominate the rhythm during developmental years, which later is succeeded by a more target-appropriate pattern once the language is mastered (Grabe, Post & Watson, 1999).

The developmental pattern discussed in the above studies illustrates an early evolution of infants' perception resulting in finer discrimination of languages according to the rhythmic pattern. The infants have an innate sensitivity to the rhythmic cues, independent from any knowledge of their native language. With the perception of rhythm cues, the children in the developmental age learn to comprehend the word boundaries in the flow of speech and acquire and assign meaning to novel words (Nazzi & Ramus, 2003). Evidence also suggests that the ability to segment the speech flow and establish word boundaries can be considered a prerequisite for successful language development (Newman, Ratner, Jusczyk, Juszcyk & Down, 2006). The strong correlation between phonological awareness in children and metrical skills as in rhythm perception, which is crucial for language acquisition, was suggested by Flaugnaco et al. (2014) in their study. Savithri, Sreedevi, Jayakumar, and Kavya (2011) did a study in typically developing Kannada speaking children (4-5y) and found that the children showed developmental trends towards the adult pattern of rhythm production (mora-timed). It was also found that boys were able to exhibit adult-like rhythm patterns by 11-12 years.

The hearing sensitivity of an individual can affect prosody because it provides necessary feedback for control over the speech. Auditory feedback affects both moment-to-moment and later control of speech, which also plays a significant role in the acquisition of language, initially with lexical stress, grammatical categories, and even emotional affect (Kuhl, 2000; Yoshinaga-Itano, 2000). van Wieringen and Wouters (2015) reported that children with cochlear implants or hearing aids show high variability in spoken language acquisition compared to their typically developing peer group. The spoken language of children with hearing impairment was found to have inappropriate pragmatics (Most et al., 2010), communication breakdowns, frequent silences with more significant duration (Tye-Murray, 2003), and markedly long speech utterances (Toe & Paatsch, 2013), all of these contributing to unnatural speech with reduced speech intelligibility.

Savithri, Jayaram, Kedarnath, and Goswami (2007) conducted a study to find the rhythm in selected Indo-Aryan and Dravidian languages. The nPVI values of the speech sample ranged between 34.09 and 57.99 with a mean of 44.96, and rPVI values ranged between 30.64 and 57.16 with a mean of 48.42. The results indicated low nPVI and high rPVI values, classifying the Malayalam language into syllable-timed rhythm category. Savithri, Ruchi, and Johnsi Rani (2008) used similar methods to study Kannada-speaking children with hearing impairment. They found that children with hearing impairment in the age range of 5-10 years had longer intervocalic and vocalic durations in comparison with their age-matched, typically developing children. The mean nPVI and rPVI values for typically developing children were 62.49 and 15.70, whereas that for the hearing-impaired group was 20.54 and 67.14. As the values were high in nPVI and low in rPVI, both groups' rhythm was classified as unclassified. The results revealed that children with and without hearing loss used a much simpler syllabic structure in comparison to the adult speech structure. Merin and Savithri (2016) studied Malayalam speaking children with hearing impairment in the age range of 3-4 years. Though Malayalam was classified as a syllable-timed language by Savithri et al. (2007) on adult native speakers of Malayalam, the speech rhythm in both children with and without hearing impairment was classified under mora-timed language. The children with hearing impairment were found to have an nPVI value of 0.223 and rPVI value of 0.293, while children with normal hearing sensitivity had nPVI of 0.184 and rPVI of 0.286. The authors emphasized the need to include intervention strategies for speech rhythm.

1.3 Studies on Speech Intonation

Typically developing children have a heightened sensitivity to variations in the suprasegmental features and, thereby, are quick in perceiving the emotions of the speaker. According to the study done by Werker and Yeung (2005), infants start picking up the acoustic cues of stress, intonation, and rhythm, to understand the meaning of the utterance, much before starting to grasp and categorize the phonemes of their particular language. These features give the infants reliable cues in comprehending the speaker's communicative intent and emotional state, which later plays a vital role in developing the social and language domains. Grossman, Striano, and Friederici (2005) studied the auditory perception of infants as young as seven months old by presenting semantically neutral words in happy, sad, and neutral voices. The emotionally loaded words evoked more excellent sensory processing when compared to neutral words. The results showed that the human brain detects the emotional content of words very early in development. Chatterjee et al. (2015) reported that in young children, the ability to identify the mood/intent of the speaker becomes a stepping stone to the acquisition of social and linguistic development. There has been a variety of studies that pointed out that the emotional content in speech has a significant role in facilitating attention (Murphy & Zajonc, 1993; Hermans, Houwer, & Eelen, 1994), cognitive performance (Bradley, Greenwald, Petry, & Lang, 1992; Kensinger & Corkin, 2003) and, memory (Kensinger & Corkin, 2003; Armony, Chochol, Fecteau, & Belin, 2007). Dupuis and Pichora-Fuller (2014) conducted a study to find if vocal emotion affects word recognition in the presence of noise. The results revealed that the presence of emotional cues could increase word recognition in both younger and older age groups.

1.3.1 Acoustic cues for Speech Intonation

The three major categories of acoustic cues that shape emotion perception are speech prosody, voice quality, and vowel articulation (Scherer, 2003). The significance of auditory cues to emotion perception was arranged in descending order by Murray and Arnott (1993) as fundamental frequency, duration, and intensity. Among these, the variability of fundamental frequency and mean fundamental frequency accounts for the perceived vocal pitch; the differences in durations at the sentence, word, and phrase levels accounts for the overall perceived duration, and the variability in intensity and mean intensity accounts for perceived vocal loudness, which all together determines the mood of the perceived speech.

Thereby when there is an emotional expression, even though voice pitch is the most dominant cue for prosody, other cues, the intensity, and duration, also change to inform the listener of the heightened emotion. Thus, even when there is a distortion in the perceived frequency cues, the listener will be able to accurately perceive the speaker's mood using the other acoustic cues.

The intonation of the Western and European languages has been extensively studied over many years. Some researchers have been curious enough to study the intonation of Indian languages using similar methods (Sethi, 1979; Manjula, 1979; Nandini, 1985; Ravishankar, 1987). Nataraja (1981) attempted to note the stress and intonation in four different South Indian languages in an emotional context. Mini (2000) studied the intonation of the Malayalam language in an emotional context. The study results showed that the perception of emotional intonation majorly depends on the variations in fundamental frequency.

1.3.2 Perception and production of intonation in individuals with hearing impairment

When there is a sensory-auditory deficit, the perception of verbal signals becomes difficult along with the auditory prosodic cues. Difficulty in the perception of the emotional content of the information conveyed can develop a lack of awareness of the social situations and deficits in social skills (Mellon, 2000). In their study, Dyck, Farrugia, Shochet, and Holmes-Brown (2004) concluded that the performance on emotion identification tasks in hearing-impaired children reflects their delayed acquisition of a broad range of language-mediated abilities. In most congenital sensory neural cases, the lower speech frequencies tend to have higher residual hearing, and most of the emotional prosodic cues are low-frequency content. However, for the more accurate perception of emotional content, the temporal fine structure characters become deficits in children with sensory neural hearing loss.

The studies comparing the perception of prosody in children with hearing impairment fitted with hearing aids and the typically developing peer group has shown a significant difference in the scores of both the group, with the children with hearing impairment showing a lower performance than their normal-hearing peer group (Most & Peled, 2007; Most & Aviner, 2009; Most & Michaelis, 2012). Many studies investigated the perception of prosody through hearing aids and have shown a mixed set of results. Jackson and Kelly (1986) found that hearing aid wearers had difficulty perceiving stress patterns in speech. Similar findings were previously reported by Stark and Levitt (1974) when they studied the reception and production of stress, intonation (yes or no questions), and location of pausal juncture in children with hearing impairment. The results showed the highest error rates in the reception and production of intonation values. Some studies that were done later report contradictory results suggesting that hearing aid users had a near-normal perception of prosodic characters like syllable stress and intonation patterns (Most, 1999). The study conducted by Most and Michaelis (2012) on the perception of the emotions through 3 modes: auditory, visual, and auditory-visual showed that in children with hearing impairment, though the overall performance was lower, the auditory-visual mode surpassed the auditory-only or visual-only modes, indicating that they relied on auditory cues for emotion perception. In their study, Boothroyd and Eran (1994) even showed that children using hearing aids had a superior perception of intonation compared to children with a cochlear implant. Despite having lower vocabulary scores, children with hearing impairment performed at the same level as typical children, according to the study of Laugen, Jacobsen, Rieffe, and Wichstrøm (2016). Ziv, Most, and Cohen (2013) also reported a similar perception of emotions by kindergarten children with hearing impairment and children with normal hearing.

Early investigations done on the intonation patterns produced by deaf children reported that the range involved in their intonation productions was 8-10 semitones less than the regular pattern (Voelker, 1935; Green, 1956). Children with hearing impairment lack a conceptual appreciation of what pitch is (Anderson, 1960) and thus have the most difficulty expressing this acoustic cue (Boothroyd,1970). Many researchers have reported the speech of children with hearing impairment to lack tonal variations or monotonous (Calvert, 1962; Martony, 1968). Sorenson (1974) analyzed the speech of deaf children and reported that they produced inappropriate or insufficient pitch change at the end of a sentence. The terminal pitch rises like that happens at the end of a question. The end of a happy emotion expression was observed to be more challenging to produce than a terminal pitch fall, as reported by Philips, Remillard, Bass, and Pronovost (1968). Abberton, Fourcin, and Hazan (1991) reported similar findings as English-speaking children with hearing aids were able to achieve falling intonation contour before rising contour. Also, Most and Frank (1994) found children with hearing impairment in the age range of 5-12 years to produce falling intonation more easily than rising.

Even when studies show that children using hearing aids could perceive intonation cues better than children with cochlear implants (Most & Peled, 2007), there is limited importance given to intonation training in speech rehabilitation of children with hearing aids. Many studies have reported the intonation of deaf individuals to play a good role in speech intelligibility. The perceived intelligibility of deaf speech was influenced majorly by intonation contours (Maassen & Povel, 1984, 1985; Metz, Schiavetti, Samar, & Sitler, 1990). Studies have reported that exposure to appropriate intervention tasks aimed at perception and production of intonation to be successful (Most & frank, 1991; Most, 2000). There is an increasing need to take up suprasegmental training of children with hearing impairment as a crucial part of the rehabilitation to increase the overall intelligibility and naturalness of speech.

Need for the study

In spite of receiving early intervention and intensive speech stimulation, the naturalness of the speech of children with hearing impairment is not that well developed. The language skills of the children with hearing impairment receive major attention while the verbal skills remain poor. Many a time, the improved articulation also does not prove to improve their spoken language comprehension. Hence, the prosodic characters of speech should be studied to improve the naturalness of speech of children with hearing impairment, which can improve spoken language comprehension.

Aim of the study

The aim of the study was to investigate the perception and production of prosody (stress, rhythm, and intonation) in children with hearing impairment.

Objectives of the study

1. To measure the perception of stress, intonation, and rhythm in children with hearing impairment and their language-age matched typically developing children.
2. To compare the perception of stress, intonation, and rhythm among children with hearing impairment and their language-age matched typically developing children.
3. To measure the production of stress, intonation, and rhythm in children with hearing impairment and their language-age matched typically developing children.
4. To compare the production of stress, intonation, and rhythm among children with hearing impairment and their language-age matched typically developing children.
5. To formulate a developmental profile for prosody in children with hearing impairment.

CHAPTER 2

METHODS

Participants

The control group of the study consisted of 30 children (16 males and 14 females) with normal hearing (CWNH) and native speakers of Malayalam. The children underwent pure tone audiometry and speech audiometry testing to confirm normal hearing sensitivity. Also, the age adequate speech and language skills were confirmed using the Modified Receptive Expressive Language Tool (M-RELT) (Deepa, Shyamala & Deepthi, 2014). The selected children were in the age range of 3.1 to 6.11 years of age (mean age = 5.4years). The children selected into the control group did not present any previous history of speech, language, or hearing deficits as ascertained by the information provided by their parents or guardians.

Thirty children (18 males and 11 females) with severe to profound sensory neural hearing impairment (CWHI) using programmable digital behind the ear hearing aids were selected for the experimental group. The children were evaluated using M-RELT to have a language age within 3 to 7 years. The authors performed aided pure tone and speech audiometry to confirm the within speech spectrum thresholds. All the CWHI group participants were using digital programmable hearing aids programmed according to DSL-IO prescriptive formula. None of the participants had any associated problems. Four participants (CWHI2, CWHI3, CWHI11, and CWHI29) were reported to have attended 1-2 years of pre-school training, and three participants (CWHI6, CWHI18, and CWHI25) were attending pre-school at the time of the study. 86.67% of the caretakers of the children with hearing impairment who reported with the child had a minimum under graduation educational qualification. The subjective factors of the children with hearing impairment like chronological age, age of hearing aid fit, duration of intensive intervention, language age, and type of communication used are listed in Table 2.1.

Table 2.1. Subject factors of the children with hearing Impairment (CWHI)

Participants	Chronological Age (in years)	Age of hearing aid fit (in years)	Duration of intensive intervention (in years)	Language age (in years)	Type of Communication
CWHI1	9.0	2.9	5.1	6.6-7.0	Oral
CWHI2	8.6	2.0	4.6	5.6-6.0	Oral
CWHI3	9.2	3.9	4.1	6.0- 6.6	Oral
CWHI4	7.1	2.6	3.9	5.6-6.0	Oral
CWHI5	4.3	1.9	2.4	3.6-4.0	Oral
CWHI6	4.4	1.6	2.8	3.6-4.0	Oral; Gestural
CWHI7	5.6	2.1	3.5	3.6-4.0	Oral
CWHI8	5.6	2.1	3.5	3.6-4.0	Oral
CWHI9	6.5	1.6	4.2	4.6-5.0	Oral
CWHI10	4.3	2.0	2.3	3.0-3.6	Oral
CWHI11	4.0	1.0	2.2	3.0-3.6	Oral; Gestural
CWHI12	4.1	2.4	1.5	3.0-3.6	Oral
CWHI13	8.6	5.3	4.2	5.0-5.6	Oral
CWHI14	5.6	2.8	2.8	5.6-6.0	Oral
CWHI15	4.8	2.4	2.1	3.6-4.0	Oral
CWHI16	10.0	1.7	1.9	6.6-7.0	Oral
CWHI17	4.6	2.4	2.0	3.6-4.0	Oral
CWHI18	5.5	2.11	3.4	3.0-3.6	Oral; Gestural
CWHI19	8.3	1.7	2.9	6.0-6.6	Oral
CWHI20	6.11	2.4	2.2	5.6-6.0	Oral; Gestural
CWHI21	4.6	2.1	2.1	3.0-3.6	Oral
CWHI22	4.5	1.5	2.10	3.0-3.6	Oral
CWHI23	5.10	2.4	2.4	4.0-4.6	Oral
CWHI24	6.2	3.1	2.10	4.6-5.0	Oral
CWHI25	5.8	1.2	3.11	4.6-5.0	Oral
CWHI26	5.5	1.8	2.5	3.0-3.6	Oral
CWHI27	6.9	2.4	2.8	4.0-4.6	Oral
CWHI28	4.3	1.8	2.3	3.6-4.0	Oral; Gestural
CWHI29	7.3	2.8	3.9	5.0-5.6	Oral; Gestural
CWHI30	8.5	2.4	5.1	6.0-6.6	Oral

Stimulus material

Twenty-four commonly occurring sentences from the kindergarten and lower primary school Malayalam textbooks were shortlisted by the authors. The selected sentences were concrete and representable in pictures with a sentence length of 3 words. Picture cards (one pair- stressed and unstressed per sentence) were custom drawn for each sentence, representing the unstressed/stressed words in the sentence. The picture representations of the stressed words were made more colorful/brighter/larger in size compared to its unstressed cognate. The pictures were validated for their representability by three typically developing Malayalam-speaking children in the age range of 3.0-6.3 years. After a clear examination of the lexical content of the sentences, ten sentences were selected as the stimuli sentences, and 14 sentences were categorized as sentences for training. The final 24 sentences and their age level distribution were validated by two other Malayalam-speaking qualified speech-language pathologists. The list of the stimuli sentences (with the stressed word underlined) is given in Table 2.2, along with the age level.

Table 2.2 List of stimuli sentences of the two age groups considered.

Sentences	Age Range (years)
1. □□□□□□□□□□□□□□□□□	
/ka:kka paṛaṉṉU vaṉṉU/	
2. □□□□□□□□□□□□□□□□□	
/kUṭṭɪ pa:l kUḍɪtʃU/	
3. □□□□□□□□□□□□□□□□□	
/tʃetʃɪ pa:ṭṛam kaḻUkɪ/	3.0 - 5.6
4. □□□□□□□□□□□□□□□	
/paʃU pUllʌ ṯɪṉṉU/	
5. □□□□□□□□□□□□□□□□□□□□	
/tʃeṭṭant$^{\varepsilon}$ baluṇ poṭṭɪ/	
6. □□□□□□□□□□□□□□□□□□□	
/atʃan kaserajɪl ɪrUṉṉU/	
7. □□□□□□□□□□□□□□□□□□□□□□□□□	

/a:ppɪḷ kaṯɪkoṇḍ^ muṛɪtʃU/	
8. □□□□□□□□ □□□□□□ □□□□□□□	
/tʃᵋḍɪkk^ vᵋḷḷam oḻɪtʃ^/	
9. □□□□ □□□□□□ □□□□	5.6 – 7.0
/valɪja ka:tt^ viʃɪ/	
10. □□□□□□□□□□□□□□□□□□	
/kUṭṭɪkaḷ kaḷɪ ṯUḍaᵑɪ/	

The 24 sentences with the two variations (stressed and unstressed) and three different emotional variations - happy, sad, and neutral; were recorded in a sound-treated room by the research officer, who was a female native speaker of Malayalam using a unidirectional microphone kept at six inches distance. Taking into consideration the decreased attention in the pediatric population, also with an added deficit of hearing impairment (Alkhamra & Abu-Dahab, 2020; Barker et al., 2009), especially while attending to a complex feature like speech prosody, the task load on the pediatric participants, the stressed-word condition of a sentence overlapped with the happy intonation and the unstressed-word condition overlapped with the neutral intonation. Acoustic analysis of the recorded stimuli was done using Praat software (Version: 5.2.01; Boersma & Weenink, 2009). Speech samples were later analyzed using the Praat software at a sampling frequency of 22 kHz with 12-bit quantization. The intensity, fundamental frequency, and average duration for the stressed and unstressed words are given in Table 2.3. The average intensity, average fundamental frequency, range of frequency, and average duration for the three emotional conditions are given in Table 2.4. From Table 2.3, it can be seen that there is a major difference between the stressed and unstressed cognate pairs in the duration measures. The stressed and unstressed expressions and the three emotional variations (happy/sad/neutral) were later evaluated perceptually by another speech-language pathologist for the accurate representation of the stress in the sentences. The recorded sentences were later normalized.

Table 2.3. The mean values of the intensity, fundamental frequency (F0), and duration (seconds) of the stressed and unstressed cognate pairs in the stimuli sentence set

Sentences	Words	Intensity (dB)	Fund. Frequency (Hz)	Duration (s)
Sentence 1	□□□□□□	81.4	285.2	0.895
	/paṛan̪n̪U/	80.7	240.9	0.580
Sentence 2	□□□	79.8	251.5	0.787
	/pa:l/	83.0	247.5	0.456
Sentence 3	□□□□□□	80.7	408.2	0.568
	/pa:t̪ram/	80.2	250.3	0.485
Sentence 4	□□□□□□	81.6	359.2	0.509
	/pUll^/	84.1	251.0	0.407
Sentence 5	□□□□	82.0	318.6	0.801
	/kasera/	81.5	255.3	0.601
Sentence 6	□□□□□	78.3	314.2	0.591
	/kat̪t̪ɪ/	78.3	262.0	0.477
Sentence 7	□□□□□□	79.4	383.6	1.10
	/poṭṭɪ/	74.9	223.6	0.504
Sentence 8	□□□□□□	82.7	307.1	0.537
	/vᵋḷḷam/	84.7	243.8	0.478
Sentence 9	□□□□	83.2	239.7	0.776
	/valɪja/	83.9	252.7	0.578
Sentence 10	□□□□□□□	80.0	353.3	0.901
	/t̪Uḍaᵑɪ/	83.5	223.6	0.556

Table 2.4. The mean values of the Average intensity, Average fundamental frequency (F0), Range (Hz), and Duration (seconds) of the stimuli sentence set used

	Avg. Intensity (dB)	Avg. F0 (Hz)	Freq. range (Hz)	Duration (s)
Happy	85.8	303.5	293.4 (26.86)	1.92
Sad	75.6	277.7	248.6 (73.35)	2.33
Neutral	80.7	246.8	161.7 (51.69)	1.56

Note. The values in parenthesis represent the standard deviation (S.D) of the frequency range of the sentences in each of the emotion conditions.

Procedure

An analytical, experimental ABA study design was adopted for this study. After selecting the participants, the study was carried out in two phases for both the control group and the experimental group. The demographic detail of the participants, along with the written consent for participation was taken from the caretakers. The participants were allotted 180 minutes of prosody perception training before evaluating their perception and production with the stimuli sentences. Phase 1 of the study included the training followed by testing in phase 2.

Phase 1: The training phase was carried out in a sound-treated audiometric room with a single room set. The auditory stimuli (14 recorded sentences categorized for training) were presented through calibrated loudspeakers at 50 dB HL at a distance of 1 meter and the azimuth of 0 degrees, routed through an audiometer along with the picture representation of each sentence. Each participant was trained and tested individually. The examiner was seated next to the participant, thus allowing the examiner to control the presentation of the auditory stimulus, picture cards, and smiley placards.

During the training phase, the examiner simultaneously presented the auditory stimulus (stressed/unstressed) and the corresponding picture card, along with the smiley emoticon placards (happy/sad/neutral) on a table in front of the child. The participants were encouraged to attend to the acoustic cues of the prosodic variations of the same sentence and choose between the picture pairs and select the appropriate smiley placard. The examiner made the participants reason out their selection of the picture by associating the stressed auditory stimulus with a bigger/brighter/colorful representation of the word. The vocal emotional variations were associated with rising, falling, and flat intonations to the happy, sad, and neutral smiley placards, respectively. A maximum of 3 sentences was selected per session (according to the child's co-operation), and during the final 10minutes, the sentences of that session were presented in randomized order. The participants were expected to point out the cards representing the stressed/unstressed word variations and the emotional variations.

The prosody production training was also carried out simultaneously, where the participants were encouraged to imitate the way of production of the recorded sentences. For stress and rhythm, production training emphasis was given on attending to the pitch, loudness, and duration of each constituent word of the recorded sentences and encouraging the participants to imitate the same model. Using the Praat software version 5.2.01 (Boersma & Weenink, 2009), children were given feedback on the pitch, loudness (online and offline), and duration parameters along with comparisons with the waveforms of the recorded sentences. The examiner also used line drawings to augment the understanding of speech intonation and encouraged the participants to imitate the same pattern. The training sessions were allotted as six - half an hour sessions, placed on either consecutive or alternate days for six days, according to the availability of the participants. The caretakers of the children were counseled not to give any kind of home training similar to the activities carried out during the sessions to maintain the homogeneity of the amount of training received by the children.

Phase 2: The testing was carried out in the same setup as in Phase 1 as a single sitting or two sittings (according to the co-operation of the participant). The testing duration varied from 30minutes to one hour, according to the participant’s response. Efforts were made to have the testing phase on the next day of the termination of the training phase; else, it was carried out on the nearest possible day (according to the availability of the participant). During Phase 2, the set of 10 recorded stimuli sentences (with the stress and intonation variations- 30sentences) were presented in randomized order, and each participant was tested individually. A maximum of two trials was given to each participant if required.

Analysis and scoring of the responses

a) Analysis of Speech Stress

Perception of Speech Stress

A forced-choice response was followed. Each child was asked to specify their response by pointing to the appropriate picture and to select the smiley placard after listening to the auditory stimuli. The correct responses were given a score of 1, and wrong responses were given a score of 0. The response of each child was tabulated for the ten stimuli sentences for each emotion. Later, to compare the stressed and unstressed perception, the tabulated values were fed into SPSS software Version 20. A normality test using Shapiro Wilk's test was performed for all data sets, which determined the use of parametric or non-parametric tests. Mann-Whitney U test was used to compare CWNH and CWHI, while paired t-test (parametric) or Wilcoxon signed-rank test was used for pre-training and post-training comparisons within groups.

Production of Speech Stress

For the acoustic analysis, the Praat software version 5.2.01 developed by Boersma and Weenink (2009) was used. The acoustic analysis was carried for the unstressed, and the identified primary stressed word in the test sentences and the fundamental frequency (F0), intensity (I0), and duration measures were extracted. The fundamental frequency (F0) and intensity (I0) values were measured directly from the pitch and intensity contours. The duration of the primary stressed syllable was taken as the time difference between the onset and offset of the intonation unit of the primary stressed word. The acoustic parameters were obtained for all the participants and were tabulated differently for CWNH and CWHI groups. To compare between the stressed-word and unstressed-word productions, the tabulated values were fed into SPSS software Version 20. A normality test using Shapiro Wilk's test was performed for all data sets, which determined the use of parametric or non-parametric tests. Independent t-test (parametric) or Mann-Whitney U test (non-parametric) was used for comparisons between CWNH and CWHI, while paired t-test (parametric) or Wilcoxon signed-rank test was used for pre-training and post-training comparisons within groups.

b) Analysis of Speech Rhythm

Perception and production of speech rhythm

The responses of the participants were recorded into an audio recorder, Olympus –WS-100, which were later fed into the Praat software version 5.2.01 developed by Boersma and Weenink (2009) for speech analysis. The recorded neutral sentence waveform responses of the participants were subjected to rhythm analysis. The pauses in the sentences were then eliminated in the waveforms so that an accurate measurement of the vocalic and inter-vocalic segment durations could be made. Each vocalic and intervocalic segment was selected individually, and the duration values were recorded in a Microsoft Office excel spread sheet. The vocalic interval duration was considered from the onset of the voicing till the offset of voicing. Durations of diphthongs, semivowels, and approximants (White & Mattys, 2007) were also considered under vocalic duration. The intervocalic duration or the consonantal duration was considered from the offset of the n^{th} vowel voicing to the onset of the $(n+1)^{th}$ vowel voicing.

The segmentation of the vocalic and inter-vocalic segments was done according to Peterson and Lehiste (1962). The vocalic and intervocalic segments were identified with amplitude differences and a break in the second formant structure. The intervocalic segments were identified using the spectrographic cues associated with the manner of articulation. Significant prolongations of the final speech segment observed in most of the speech samples were excluded in the analysis to maintain uniformity throughout the data. Also, aspiration following the release of stop consonant was considered within the consonantal duration (White & Mattys, 2007).

Once the vocalic (VD) and intervocalic durations (IVD) were calculated, the Normalised Pairwise Variability Index-vocalic (nPVI-V) and Raw Pairwise Variability Index-consonantal (rPVI-C) scores for the duration were calculated using the formulae given by Low, Grabe, and Nolan (2000):

$$nPVI = 100 \times \frac{\left(\sum_{k=1}^{m-1} \left| (d_k - d_{k+1})/((d_k + d_{k+1})/2) \right| \right)}{(m-1)}$$

$$rPVI = \left(\sum_{k=1}^{m-1} \left| d_k - d_{k+1} \right| \right) / (m-1)$$

Where,

m—number of intervals in an utterance for which PVI is calculated,

d—duration of the kth interval.

The PVI value above 50 was considered high, and below 50 was considered low (Merin, 2016). The high and low PVI rhythm class of each age group was categorized into stress-timed, syllable-timed, and mora-timed rhythm classes, as shown in Table 2.5.

Table 2.5. Rhythm class classification according to PVI values.

	Stress-timed	Syllable-timed	Mora-timed
nPVI	High	Low	Low
rPVI	High	High	Low

Note. nPVI- normalized Pairwise Variability Index (nPVI); rPVI- raw Pairwise Variability Index.

The participants who were able to replicate the rhythm pattern of the stimuli were considered to have perceived the rhythm pattern of the stimuli correctly. The mean PVI values of the age groups were considered to classify the rhythm pattern into stress-timed, syllable-timed, and mora-timed languages.

Statistical analysis of the data was carried out in the commercially available SPSS (version 20) software. Normal distribution of the nPVI-vocalic and rPVI-intervocalic was determined. Descriptive and inferential statistics were carried out using the statistical package SPSS version 20. Descriptive statistics were done to determine the mean, median, and standard deviation. To find if there was a significant difference between the CWNH and CWHI groups across the age groups, Mann Whitney or Independent t test and Wilcoxon Signed Rank test o Paired sample t test statistical tests were carried out.

c) Analysis of Speech Intonation

The recorded responses of the participants were analyzed for the different intonation units constituting the sentences and the primary stressed intonation unit among the identified ones. An intonation unit was defined as a sequence of syllables/words combined under a single coherent intonation contour (Chafe, 1987). The intonation units in the responses were restricted according to the criteria given by Cruttenden (1997), which were: the presence of at least one stressed syllable, significant pause between intonation units, final phrase lengthening, anacrusis, and pitch reset. The intonation units were identified independently by one more Speech-Language Pathologist (SLP) experienced in suprasegmental analysis. An item-by-item inter-judge reliability coefficient ‘alpha’ for identification of intonation units was found to be 0.9275. The intonation unit of the word with primary stress in the sentence was identified and analyzed for acoustic parameters.

Perception of Speech Intonation

A forced-choice response was followed. Each child was asked to specify their response by pointing to the appropriate smiley after listening to the stimuli pairs. The correct responses were given a score of 1, and the responses of each child were tabulated for the ten stimuli for each emotion. The response of all the participants was tabulated and was subjected to suitable statistical analyses.

Production of Speech Intonation

For the analysis of the intonations produced by the participants, the samples were analyzed in the F0 analysis module of the PRAAT software version 5.2.01 (Boersma & Weenink, 2009). The pitch extracting algorithm of Praat software extracted the pitch contours of all utterances. The pitch intonation units of the pitch contours were identified, and the final intonation unit of the pitch contour was subjected to temporal and F0 measurements. The reliability of temporal and F0 parameters measurements was measured independently on approximately 10% of the data by another Speech-Language Pathologist familiar with the contour identification and intonation measurements. The measures of onset frequency, offset frequency, and duration of the transition were measured and tabulated. The tabulated values were fed into SPSS version 20 and evaluated for finding significant differences between CWNH and CWHI groups using statistical tests.

Statistical Analyses

The mean and standard deviation of the parameters of interest were calculated using descriptive statistics. A normality test using Shapiro Wilk's test was performed for all data sets, which determined the use of parametric or non-parametric tests. Independent t-test (parametric) or Mann-Whitney U test (non-parametric) was used for comparisons between CWNH and CWHI, while paired t-test (parametric) or Wilcoxon signed-rank test was used for pre-training and post-training comparisons within groups.

CHAPTER-3

RESULTS

The results of the study are presented as perception and production of each of the groups; children with normal hearing (CWNH) and children with hearing impairment (CWHI) and comparison of the perception and production of both the groups on each prosody parameter: Stress, Rhythm, and Intonation. The objectives of the study will be discussed together under each section, which was: -

1. To measure the perception of stress, intonation, and rhythm in children with hearing impairment and their language-age matched typically developing children.
2. To compare the perception of stress, intonation, and rhythm among children with hearing impairment and their language-age matched typically developing children.
3. To measure the production of stress, intonation, and rhythm in children with hearing impairment and their language-age matched typically developing children.
4. To compare the production of stress, intonation, and rhythm among children with hearing impairment and their language-age matched typically developing children.
5. To formulate a developmental profile for prosody in children with hearing impairment.

3.1 Perception and comparison of speech prosody in CWNH and CWHI

3.1.1 Perception of speech stress

The first objective of the study was to measure the stress perception scores of children with normal hearing (CWNH) and children with hearing impairment (CWHI). Shapiro Wilk's test of normality was done to determine the normal distribution of the scores obtained by the participants. The results showed a significance of $p > 0.05$, indicating that the data was normally distributed. The mean and standard deviations of the scores obtained by both the groups, CWNH and CWHI, under the two prosodic conditions are given in Table 3.1. From Table 3.1, it can be seen that both the groups performed better on stressed-word conditions than on unstressed.

Table 3.1 The mean and standard deviation of the perception scores obtained by the participants under stressed-word and unstressed-word conditions.

	Stressed-word		Unstress-word	
	Mean (Max = 10)	S.D	Mean (Max = 10)	S.D
CWNH	8.5	0.23	7.2	0.22
CWHI	7.5	0.26	6.1	0.22

Note. CWNH- Children with Normal Hearing; CWHI – Children with Hearing Impairment. S.D – Standard Deviation; Max – Maximum score

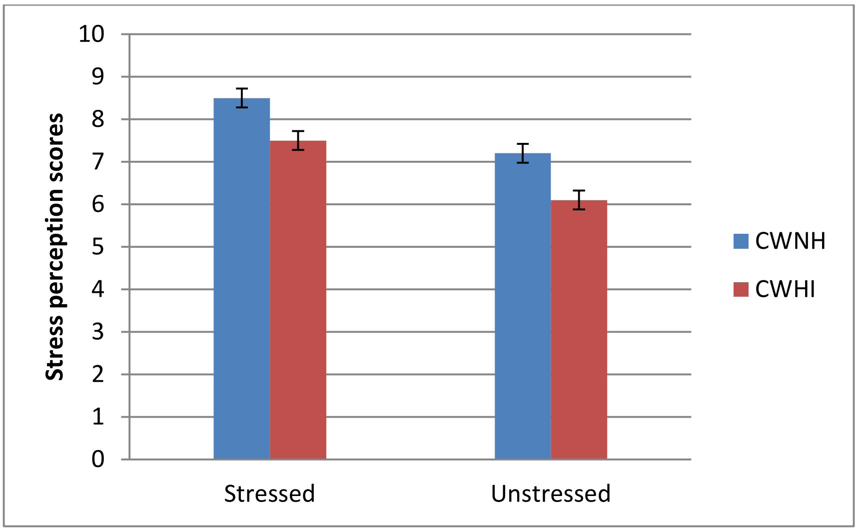

Figure 3.1. The mean and standard deviation of the perception scores of control group (CWNH) and experimental group (CWHI) under the conditions: stressed-word and unstressed-word.

Note. CWNH- Children with Normal Hearing; CWHI – Children with Hearing Impairment. Error bars represent +/- Standard Deviation (S.D).

As the data were normally distributed, paired t-test was done to compare the stressed-unstressed perception scores of CWNH and CWHI. The results showed that the CWNH perception scores of stressed (M= 8.56, S.D = 1.30) and unstressed (M= 7.26, S.D = 1.25) words were significantly different; t (29) = 5.76, p = 0.00. Interestingly, the paired t-test done on stressed (M= 7.53, S.D = 1.45) and unstressed (M= 6.10, S.D = 1.24) perception scores of CWHI group also showed a significant difference; t (29) = 6.91, p = 0.00.

The second objective of the study was to compare the perception scores between CWNH and CWHI. Mann-Whitney U test was performed on the perception scores under stressed-word and unstressed-word conditions. The results showed that there was a significant difference between groups under both stressed-word (│Z│ = 2.68; p = 0.007; r = 0.04) and unstressed-word (│Z│ = 3.21; p = 0.01; r = 0.05) conditions. Mann-Whitney U test was done across age groups to compare the difference in perception of stress between CWNH and CWHI groups. No significant differences were found in any age group under both stressed and unstressed-word conditions (Table 3.2 & Table 3.3).

Table 3.2 Results of Mann-Whitney U Test on perception scores of CWNH-CWHI under stressed-word condition.

	3.1-3.11	4.0-4.11	5.0-5.11	6.0-6.11
│Z│ value	1.20	0.113	1.20	0.703
p value	0.229	0.910	0.905	0.482

Note. p=≤0.05

Table 3.3 Results of Mann-Whitney U Test on perception scores of CWNH-CWHI under unstressed-word condition.

	3.1-3.11	4.0-4.11	5.0-5.11	6.0-6.11
│Z│ value	1.85	0.219	0.113	1.07
p value	0.064	0.827	0.910	0.285

Note. p=≤0.05

Pearson's correlation coefficient was done to check for any correlation that could exist between language age and stress perception scores in the CWHI group to fulfill the fifth objective of the study. The results revealed a positive correlation (r = 0.98; p = 0.00) between the stress perception scores and the language age of the participants of CWHI. This indicates that as the language age increases, there is an evident increase in stress perception scores. Figure 3.2 represents the scatter plot showing a positive correlation between the language age and the stress perception score. A clear developmental trend with respect to language age in the perception of stress is seen in this scatter plot.

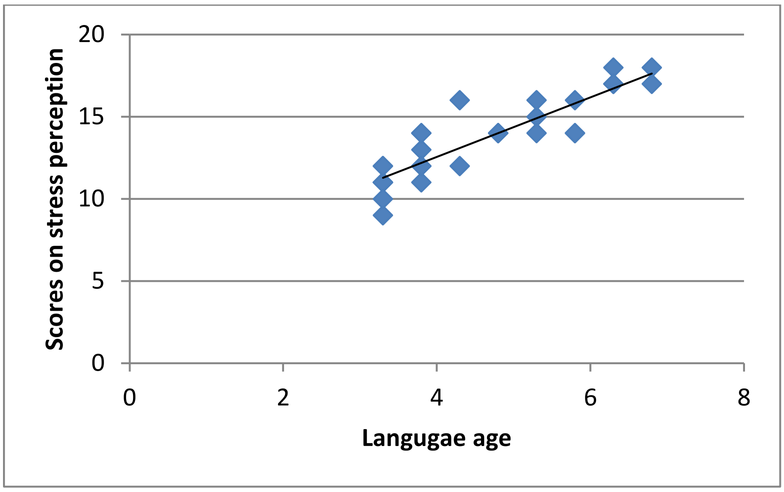

Figure 3.2 Scatter plot indicating a correlation between language age and stress perception scores.

3.1.2 Perception of speech rhythm

It was observed that the majority of the participants in the CWNH group, except for the age group of 4.0-4.11, were able to replicate the syllable-timed rhythm pattern similar to the stimuli. The rhythm pattern perceived by the participants in the age group of 3.0-6.0 years was deviant from the rhythm pattern of the stimuli. Interestingly, the eldest age group, 6.0-6.11, was able to perceive and replicate the rhythm pattern of the stimuli. The mean nPVI and rPVI values (Table 3.13) were observed to be prolonged in both CWNH and CWHI groups compared to the adult pattern.

3.1.3 Perception of intonation

The first and second objectives of the study were to measure the intonation perception scores of children with normal hearing (CWNH) and children with hearing impairment (CWHI) and compare the same. The vocal emotion perception scores of CWNH and CWHI are represented as confusions percent in Tables 3.4 & 3.5. From the percent values, it can be seen that both the groups were able to perceive the happy emotion the best (CWNH=91.3; CWHI=84.6). The perception of sad emotion (CWNH=83.3; CWHI=63.3) was better than the neutral condition (CWNH=69.3; CWHI=48.6). Both the groups exhibited an obvious confusion between the sad and neutral emotion conditions.

Table 3.4. Confusion matrix representing the responses of CWNH (In percentage)

	HAPPY	SAD	NEUTRAL
HAPPY	91.3	2	6.6
SAD	5.3	83.3	11.3
NEUTRAL	3.3	27.3	69.3

Table 3.5. Confusion matrix representing the responses of CWHI (In percentage)

	HAPPY	SAD	NEUTRAL
HAPPY	84.6	6.6	8.6
SAD	2.6	63.3	34
NEUTRAL	18.6	32.6	48.6

Shapiro Wilk’s test of normality was carried out to determine the normality of the perception scores obtained from the sample population under the three conditions happy, sad and neutral concerning the independent variable of the presence of hearing loss. The results revealed that the data under the conditions of sad and neutral emotional prosodies were normally distributed ($p>0.05$), while the scores obtained by both the groups under the happy emotional prosody were not normally distributed ($p<0.05$). The mean and standard deviations of the perception scores of the CWNH and CWHI are mentioned in Table 3.6 and Figures 3.3 & 3.4. Both the control and experimental groups followed the same trend of scores, obtaining better scores in happy emotion perception, followed by sad emotion, and scored the least in neutral emotion perception.

Table 3.6. The median, mean and standard deviation of the perception of emotions

GROUPS	EMOTIONS						
	HAPPY			SAD		NEUTRAL	
	Median	Minimum	Maximum	Mean	S.D	Mean	S.D
CWNH	10.0	7.0	10.0	8.33	1.49	6.93	1.57
CWHI	8.0	7.0	10.0	6.33	1.58	4.86	2.23

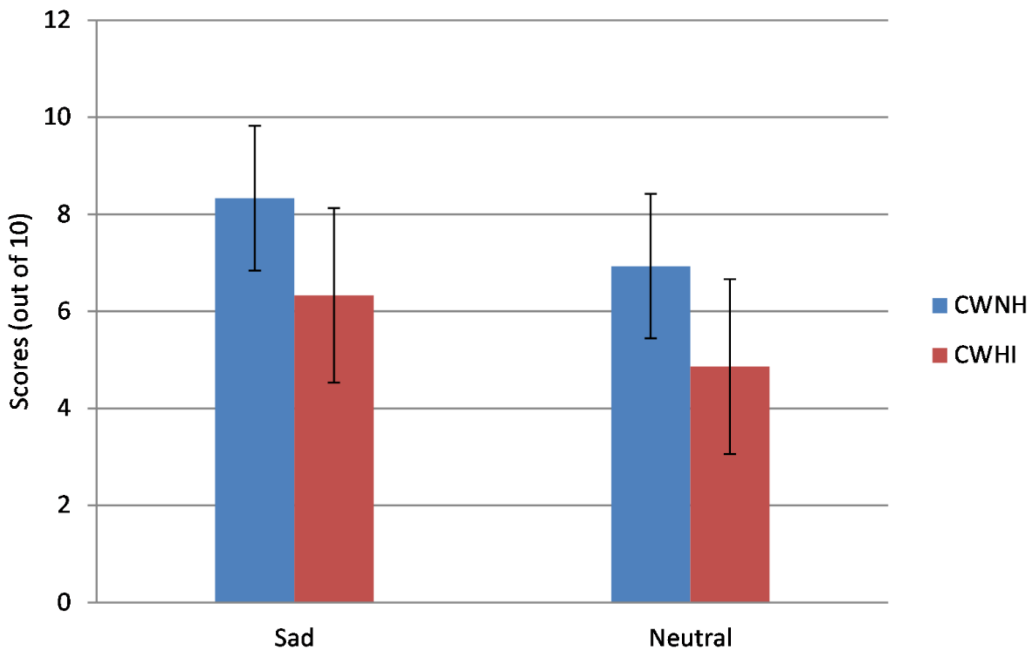

Figure 3.3. The mean and standard deviation of emotion perception scores of control group (CWNH) and experimental group (CWHI) under the emotion conditions: sad and neutral.

Note: CWNH- Children with Normal Hearing; CWHI – Children with Hearing Impairment. Error bars represent +/- Standard Deviation (S.D).

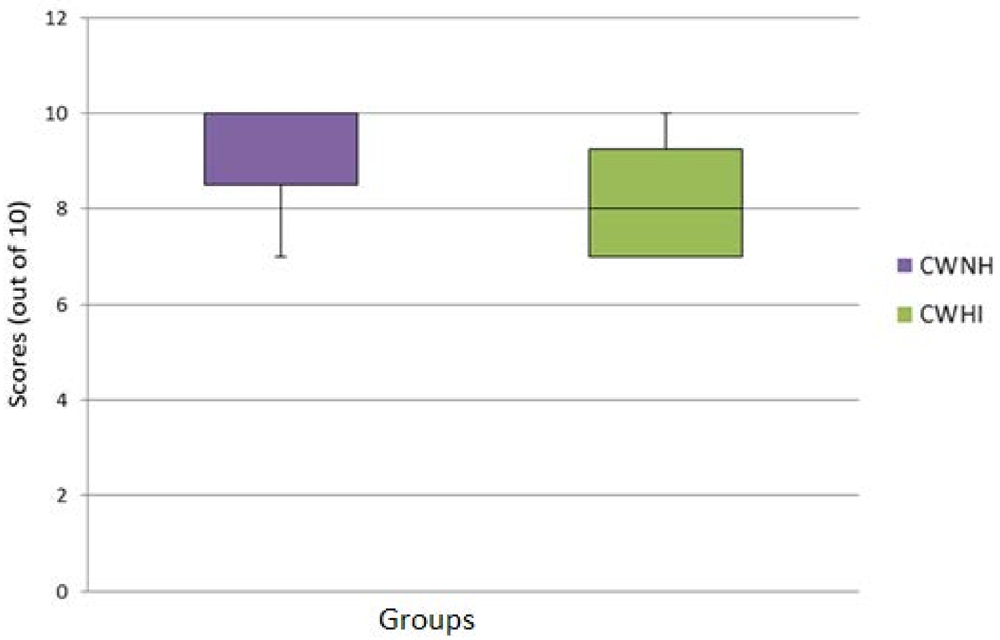

Figure 3.4. Box plot showing the upper extreme, upper quartile, median, lower quartile and lower limit for the scores achieved by CWNH and CWHI on emotion perception.

To compare the perception scores of CWNH and CWHI, Wilcoxon Signed Rank Test was used, and the perception of happy–emotional prosody between CWNH and CWHI groups revealed no significant difference (Z= 1.53; p=0.125) between the groups. Further, the paired sample t-test was used to compare the emotional prosody perception scores of CWNH and CWHI in the conditions of sad and neutral prosody. The results showed that the perception scores were significantly different between the groups CWNH (Mean=8.33; S.D=1.49) and CWHI (Mean=6.33; S.D=1.58) in sad emotional condition; t(14)=3.05, p=0.009. Also, under the neutral emotion condition, there was a significant difference between the groups CWNH (Mean=6.93; S.D=1.57) and CWHI (Mean=4.86; S.D=2.23); t(14)=2.50; p=0.025. The examination of the emotion scores revealed that the CWHI showed significantly poorer performance in the sad and neutral emotion conditions.

Pearson Correlation coefficient was found to draw relations between the language age and the total perception scores of CWHI participants. The result of the Pearson correlation revealed a significant positive correlation between the perception scores and the language

age of the CWHI participants (r=0.86; p=0.000), indicating an increase in the perception scores with the increase in language age (Figure 3.5).

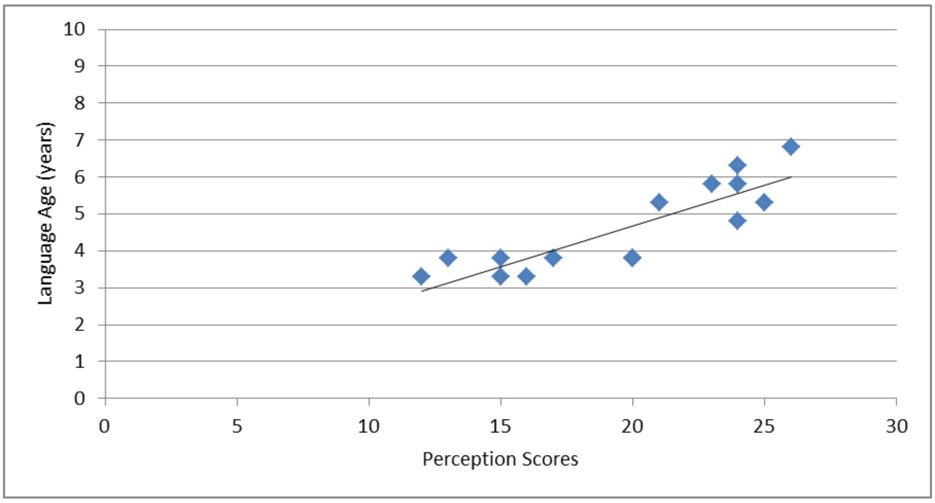

Figure 3.5 Scatter plot indicating a correlation between language age and intonation perception scores.

3.2 Production and comparison of speech prosody in CWNH and CWHI

3.2.1 Production of speech stress

According to the third objective of the study, dominant acoustic cues for stress which include fundamental frequency (F0), intensity (I0), and duration was measured from the responses of the participants of CWNH and CWHI groups. Further, these parameters were statistically analyzed across the age groups (3.1 -6.11 years) to complete the fourth objective of the study, which was to compare the production of prosody parameters between the CWNH and CWHI groups.

a) *Fundamental frequency (F0) of the primary stressed word:* The mean and standard deviation of the F0 of the stressed word produced by CWNH and CWHI in different age groups are given in Table 3.7. Mann-Whitney U test was done to find the significant difference between F0 measures of CWNH and CWHI. There was no significant difference found between the F0 measures of any of the age groups of CWNH and CWHI groups (Table 3.8).

Table 3.7 The mean and standard deviation (S.D) of the fundamental frequency of CWNH and CWHI group.

Age groups (years)	N		F0 CWNH		F0 CWHI	
	CWNH	CWHI	Mean	S.D	Mean	S.D
3.1-3.11	5	15	293.91	46.23	308.84	48.36
4.0-4.11	5	5	331.59	12.93	334.28	20.11
5.0-5.11	5	5	321.69	43.96	311.34	4.50
6.0-6.11	15	5	299.57	38.67	274.50	32.51

Table 3.8 Results of Mann-Whitney U test on Fundamental frequency measures of the CWNH and CWHI groups.

	3.1-3.11	4.0-4.11	5.0-5.11	6.0-6.11
$\lvert Z \rvert$ value	0.918	0.184	0.554	1.12
p value	0.359	0.854	0.580	0.263

Note. p=≤0.05

b) *Mean intensity (MI) of primary stressed word*

The mean and standard deviation of the F0 of the stressed word produced by CWNH and CWHI in different age groups are given in Table 3.9. Mann-Whitney U test was done to find the significant difference in mean intensity between CWNH and CWHI. There was a significant difference found in MI for all age groups of CWNH and CWHI groups, except for the age group of 6.0-6.11 years (Table 3.10).

Table 3.9 The mean and standard deviation (S.D) of mean intensity of CWNH and CWHI group.

Age groups (years)	N		MI CWNH		MI CWHI	
	CWNH	CWHI	Mean	S.D	Mean	S.D
3.1-3.11	5	15	63.46	5.36	60.70	7.54
4.0-4.11	5	5	63.25	14.78	70.93	15.69
5.0-5.11	5	5	65.01	2.10	67.27	2.10
6.0-6.11	15	5	64.40	2.11	68.39	2.41

Table 3.10. Results of Mann-Whitney U test on stress intensity measures of the CWNH and CWHI groups.

	3.1-3.11	4.0-4.11	5.0-5.11	6.0-6.11
$\lvert Z \rvert$ value	2.07	2.65	2.79	1.86
p value	0.038	0.008	0.005	0.062
r value	0.034	0.044	0.046	-

Note. p=≤0.05

c) *Duration of primary stressed word*

The mean and standard deviation of the duration (in seconds) of the stressed word produced by CWNH and CWHI in different age groups are given in Table 3.11. A significant difference in duration (second) was found between the age groups 3.1-3.11 years and 4.0-4.11 years of CWNH and CWHI groups. At the same time, there was no significant difference found in the age group of 5.0-5.11 years and 6.0-6.11years (Table 3.12).

Table 3.11. The mean and standard deviation (S.D) of duration in CWNH and CWHI group.

Age groups (years)	N		Duration (CWNH) (seconds)		Duration (CWHI) (seconds)	
	CWNH	CWHI	Mean	S.D	Mean	S.D
3.1-3.11	5	15	0.668	0.123	0.903	0.160
4.0-4.11	5	5	0.711	0.134	0.925	0.104
5.0-5.11	5	5	0.632	0.085	0.710	0.062
6.0-6.11	15	5	0.671	0.098	0.772	0.461

Table 3.12. Results of Mann-Whitney U test on duration measures of the CWNH and CWHI groups.

	3.1-3.11	4.0-4.11	5.0-5.11	6.0-6.11
$\lvert Z \rvert$ value	3.31	2.63	1.86	1.86
p value	0.001	0.008	0.062	0.062
r value	0.05	0.043	-	-

Note. p=≤0.05

3.2.2. Production of Speech Rhythm

According to the third and fourth objectives of the study, the rhythm parameters were measured from the responses of the participants of CWNH and CWHI groups and compared across the participants. Further, these parameters were statistically analyzed across the age groups (3.1 -6.11 years).

The Mann-Whitney U test results showed an overall significant difference between the PVI scores of the CWNH and CWHI groups (Z =2.15; p<0.05). The Mann-Whitney U test was also done across all the four different age groups (3.1-3.11; 4.0-4.11; 5.0-5.11; 6.0-6.11) of CWNH and CWHI, which revealed the groups differed significantly across all age groups.

Descriptive statistics were done to calculate the mean, median, and standard deviation of the nPVI and rPVI values, using SPSS software (Version 20). The mean values of CWHI were observed to be higher than the CWNH in all the age groups except for the oldest age group of 6.0-6.11 years, where the PVI values were almost similar to the CWNH values. It was also noted that the standard deviation was higher for the CWHI group, indicating that the group participants showed greater variability. The PVI values measured for the groups are tabulated in Table 3.13.

Table 3.13. The mean, median, and standard deviation of nPVI and rPVI values for the CWNH and CWHI across the age group.

Groups	PVI values		Age groups			
			3.1-3.11	4.0-4.11	5.0-5.11	6.0-6.11
CWNH	nPVI	Mean	38.89	59.99	40.27	49.2
		Median	39.46	59.14	40.36	49.27
		S.D	1.92	1.93	1.93	2.06
	rPVI	Mean	100.23	89.11	59.68	80.23
		Median	99.65	89.27	60.41	80.13
		S.D	1.45	1.56	1.56	2.15
CWHI	nPVI	Mean	52.92	64.66	57.01	38.74
		Median	53.01	65.24	57.32	38.46
		S.D	2.02	1.54	1.66	2.03
	rPVI	Mean	111.45	140.73	177.11	75.63
		Median	112.46	141.75	176.82	74.47
		S.D	3.67	2.97	4.36	4.97

Based on the PVI values, the rhythm patterns for the different age groups can be categorized as shown in Table 3.14.

Table 3.14. Rhythm classes of CWNH and CWHI groups, according to the PVI values.

		Age Groups			
		3.1-3.11	4.0-4.11	5.0-5.11	6.0-6.11
CWNH	nPVI	low	high	low	low
	rPVI	high	low	high	high
	Rhythm Class	Syllable-timed	Unclassified	Syllable-timed	Syllable-timed
CWHI	nPVI	high	high	high	low
	rPVI	high	high	high	high
	Rhythm Class	Stress-timed	Stress-timed	Stress-timed	Syllable-timed

3.2.3. Production of Intonation

According to the third and fourth objectives of the study, the intonation parameters were measured from the responses of the participants of CWNH and CWHI groups and compared. Further, these parameters were statistically analyzed across the age groups (3.1 -6.11 years) to complete the fourth objective of the study, which was to compare the production of prosody parameters between the CWNH and CWHI groups. The happy, sad, and neutral emotional intonation contours produced by the participants were assessed for their onset frequency, offset frequency, difference between the onset and offset frequencies, and the terminal intonation contours.

a) Happy emotion

Onset frequency of the happy intonation contours.

The descriptive statistics module of SPSS software version 20 was used to calculate the mean and standard deviation of onset frequency for each age group and is given in Table 3.15. The onset frequency of both the groups across all the age groups was observed to be similar. However, the standard deviation of the onset frequency of the CWNH group was higher when compared to the CWHI group.

Table 3.15. Mean and standard deviation of onset frequency of happy intonation contour.

	Age groups							
	3.1-3.11		4.0-4.11		5.0-5.11		6.0-6.11	
	Mean	S.D	Mean	S.D	Mean	S.D	Mean	S.D
CWNH	313.56	61.5	306.89	52.3	286.75	29.67	297.19	36.28
CWHI	320.33	36.69	291.02	10.86	317.55	15.91	300.86	13.70

The Mann-Whitney U test was done to find the significance of the difference between CWNH and CWHI groups' production of onset frequency of the happy intonation contour (Table 3.16). The results revealed no significant difference between the CWNH and CWHI groups, except for the age group of 4.0-4.11 years.

Table 3.16. Results of Mann-Whitney U Test on CWNH-CWHI onset frequency under happy intonation condition

	Age groups			
	3.1-3.11	4.0-4.11	5.0-5.11	6.0-6.11
$\lvert Z \rvert$ value	0.663	2.44	1.38	0.484
p value	0.507	0.014	0.167	0.628
r value	-	0.244	-	-

Offset frequency of the happy intonation contour

The mean and standard deviation of the offset frequency produced by the participants across age groups is given in Table 3.17. The mean offset frequency of both groups was similar, and CWHI presented a slightly reduced standard deviation than the other group.

From Table 2.1, it could be observed that 77.78% of CWHI in this age range was assessed to have a language age of 3.0-3.6 years.

Table 3.17. Mean and standard deviation of offset frequency of happy intonation contour

	Age groups							
	3.1-3.11		4.0-4.11		5.0-5.11		6.0-6.11	
	Mean	S.D	Mean	S.D	Mean	S.D	Mean	S.D
CWNH	372.23	42.85	347.51	45.13	378.05	13.63	286.03	42.53
CWHI	373.32	35.15	343.64	5.16	325.34	10.36	290.30	47.47

Mann-Whitney U test revealed no significant difference found offset frequency productions of CWNH and CWHI for the youngest (3.1-3.11) and eldest (6.1-6.11) age groups; while there was a significant difference found between the groups for the age groups 4.0-4.11 and 5.0-5.11 years (Table 3.18).

Table 3.18. Results of Mann-Whitney U Test on CWNH-CWHI offset frequency under happy intonation condition.

	3.1-3.11	4.0-4.11	5.0-5.11	6.0-6.11
$\|Z\|$ value	0.133	2.44	2.66	0.395
p value	0.894	0.016	0.008	0.692
r value	-	0.24	0.26	-

Note. p=≤0.05

Terminal contour of the happy emotion

Comparing the terminal contours of CWNH and CWHI, it was found that the majority of the CWNH participants were able to produce the rising contour. In contrast, it was difficult for CWHI participants as most of them failed to produce a rising pattern. The accuracy of the contour improved with age in both groups. The details of the percentage of participants who produced the correct rising terminal contour for happy emotion are given in Table 3.19.

Table 3.19. Percentage of participants who produced the rising intonation contour

	Age groups			
	3.1-3.11	4.0-4.11	5.0-5.11	6.0-6.11
CWNH	60	100	100	100
CWHI	26.66	20	40	40

b) Sad emotion

Onset frequency of the sad intonation contours.

The descriptive statistics module of SPSS software version 20 was used to calculate the mean and standard deviation of onset frequency for each age group and is given in Table 3.20. The mean and standard deviation of the onset frequencies of both groups were found to be similar.

Table 3.20. Mean and standard deviation of onset frequency of sad intonation contour.

	Age groups							
	3.1-3.11		4.0-4.11		5.0-5.11		6.0-6.11	
	Mean	S.D	Mean	S.D	Mean	S.D	Mean	S.D
CWNH	282.14	87.84	329.09	64.24	269.79	53.48	274.16	39.54
CWHI	322.48	37.46	269.72	23.45	305.09	15.47	283.79	26.24

The Mann-Whitney U test was done to find the significance of the difference between CWNH and CWHI groups' production of onset frequency of the sad intonation contour (Table 3.23). The results revealed no significant difference between the CWNH and CWHI groups, except for the age group of 4.0-4.11 years (Table 3.21).

Table 3.21 Results of Mann-Whitney U Test on CWNH-CWHI onset frequency under sad intonation condition

	3.1-3.11	4.0-4.11	5.0-5.11	6.0-6.11
$\lvert Z \rvert$ value	0.663	2.44	1.38	0.484
p value	0.507	0.014	0.167	0.628
r value	-	0.24	-	-

Note. p=≤0.05

Offset frequency of the sad intonation contour

The mean and standard deviation of the offset frequency produced by the participants across age groups is given in Table 3.22. The values were observed to be similar for both groups.

Table 3.22 Mean and standard deviation of offset frequency of sad intonation contour

	Age groups							
	3.1-3.11		4.0-4.11		5.0-5.11		6.0-6.11	
	Mean	S.D	Mean	S.D	Mean	S.D	Mean	S.D
CWNH	219.99	74.84	292.30	61.28	296.33	54.76	254.31	28.34
CWHI	322.48	37.46	296.33	32.15	314.72	15.04	255.72	49.68

There was a significant difference found for the offset frequency between the groups in the age groups 4.0-4.11 and 5.0-5.11 years under sad intonation condition, while there was no significant difference found between the age groups 3.1-3.11 and 6.0-6.11 years (Table 3.23).

Table 3.23 Results of Mann-Whitney U Test on CWNH-CWHI offset frequency under sad intonation condition

	3.1-3.11	4.0-4.11	5.0-5.11	6.0-6.11
$\lvert Z \rvert$ value	0.133	2.44	2.66	0.395
p value	0.894	0.016	0.008	0.692
r value	-	0.24	0.26	-

Note. p=≤0.05

Terminal contour of the sad emotion

Comparing the terminal contours of CWNH and CWHI, it was found that the majority of the CWNH participants were able to produce the falling contour with better accuracy than the CWHI group. The percentage of correct contour production was observed to improve with age in both groups. The details of the percentage of participants who produced the correct falling terminal contour for sad emotion are given in Table 3.24.

Table 3.24 Percentage of participants who produced the falling intonation contour

	Age groups			
	3.1-3.11	4.0-4.11	5.0-5.11	6.0-6.11
CWNH	80	80	100	100
CWHI	13.33	40	40	80

c) Neutral emotion

Onset frequency of the neutral intonation contours.

The descriptive statistics module of SPSS software version 20 was used to calculate the mean and standard deviation of onset frequency for each age group and is given in Table 3.25. The CWNH group was found to have a higher mean onset frequency in two age groups, 4.0-4.11 and 5.0-5.11.

Table 3.25. Mean and standard deviation of onset frequency of neutral intonation contour

	Age groups							
	3.1-3.11		4.0-4.11		5.0-5.11		6.0-6.11	
	Mean	S.D	Mean	S.D	Mean	S.D	Mean	S.D
CWNH	314.46	29.01	357.52	20.41	311.88	23.02	274.16	25.19
CWHI	326.08	41.69	268.85	35.46	310.62	7.07	280.89	25.19

The Mann-Whitney U test was done to find the significance of the difference between CWNH and CWHI groups' production of onset frequency of the neutral intonation contour. The results revealed no significant difference between the CWNH and CWHI groups, except for the age group of 4.0-4.11 years (Table 3.26).

Table 3.26 Results of Mann-Whitney U Test on CWNH-CWHI mean onset frequency under neutral intonation condition

	3.1-3.11	4.0-4.11	5.0-5.11	6.0-6.11
$\lvert Z \rvert$ value	0.921	2.887	0.532	0.573
p value	0.353	0.004	0.595	0.566

Note. p=≤0.05

Offset frequency of the neutral intonation contour

The mean and standard deviation of the offset frequency produced by the participants across age groups is given in Table 3.27. The mean and standard deviation of the neutral intonation was observed to be similar in both groups.

Table 3.27 Mean and standard deviation of mean offset frequency of neutral intonation contour

	Age groups							
	3.1-3.11		4.0-4.11		5.0-5.11		6.0-6.11	
	Mean	S.D	Mean	S.D	Mean	S.D	Mean	S.D
CWNH	285.48	35.84	299.92	21.49	291.88	20.73	254.31	28.34

CWHI	333.81	28.83	296.29	19.25	300.63	19.71	253.28	32.44

There was a significant difference found between offset frequency productions of CWNH and CWHI under neutral intonation conditions for the age group 3.1-3.11 years; while there was no significant difference found between the groups across the age groups of 4.0-4.11; 5.0-5.11, and 6.0-6.11 years (Table 3.28).

Table 3.28. Results of Mann-Whitney U Test on CWNH-CWHI mean offset frequency under happy intonation condition

	Age groups			
	3.1-3.11	4.0-4.11	5.0-5.11	6.0-6.11
$\lvert Z \rvert$ value	2.610	1.732	0.106	0.044
p value	0.009	0.083	0.915	0.965
r value	0.130	-	-	-

Note. p=≤0.05

Terminal contour of the neutral emotion

Both CWNH and CWHI groups could produce flat intonation contours accurately, except for the 3.1-3.11 years age group of CWNH. The details of the percentage of participants who produced the correct flat contour for neutral emotion are given in Table 3.29.

Table 3.29 Percentage of flat intonation contour produced by the groups across the age groups

	Age groups			
	3.1-3.11	4.0-4.11	5.0-5.11	6.0-6.11
CWNH	66.66	100	100	100
CWHI	100	100	100	100

CHAPTER-4

DISCUSSION

The results of the study will be discussed under the headings of perception and production of prosody with subsections as speech stress, speech rhythm, and speech intonation, covering the first four objectives:

1. To measure the perception of stress, intonation, and rhythm in children with hearing impairment and their language-age matched typically developing children.
2. To compare the perception stress, intonation, and rhythm among children with hearing impairment and their language-age matched typically developing children.
3. To measure the production of stress, intonation, and rhythm in children with hearing impairment and their language-age matched typically developing children.
4. To compare the production of stress, intonation, and rhythm among children with hearing impairment and their language-age matched typically developing children.
5. To formulate a developmental profile for prosody in children with hearing impairment.

4.1 Perception of speech prosody in children with hearing impairment

4.1.1 Perception of speech stress.

The first and second objectives of the study were to measure the perception and compare results across CWNH and CWHI participants. It was observed that both the groups' CWNH and CWHI scores were significantly different in both stressed-word and unstressed-word conditions. The mean scores revealed that the perception scores of CWHI participants were lower than the participants of CWNH group. This is in consensus with the numerous studies that report individuals with hearing impairment to have a poorer perception of stress acoustic cues (Jackson & Kelly, 1991). The CWHI has limited access to the spectral cues that account for the perception of speech stress, resulting in their poor performance on speech stress detection tasks. However, the literature suggests that among the different amplification devices, hearing aids can detect and discriminate the time energy envelope of speech signals, which carries information on suprasegmental cues (Erber, 1979; Engen, Engen, Clarkson & Blackwell, 1983). Kong, Stickney, and Zeng (2005) reported that hearing aids enable low-frequency acoustic information, which cannot be represented in electrical stimulations like Cochlear Implants (CI). Henry and Turner (2003) reported that hearing aids could deliver finer spectral and temporal pitch cues, which were not resolved by CIs. Most and Peled (2007) found that children using hearing aids had better suprasegmental perception scores than cochlear implanted children with hearing impairment. Even when abundant literature is available on the hearing aids to process the suprasegmental information, the performance of children using hearing aids with aided hearing thresholds within the speech spectrum is observed to be poor on speech stress and other suprasegmental perception tasks.

Both the groups CWHI and CWNH groups were found to have a significant difference in their scores of perceptions stressed-word and unstressed-word tasks. Interestingly, participants of the CWHI group were able to auditorily perceive and differentiate between the stressed and unstressed-word acoustic cues. It was noted that the participants in both groups had better scores in stressed-word conditions than the unstressed-word conditions (Table 1.1). The acoustic emphasis in intensity, frequency, and duration (Bloomfield, 1933; Trager & Smith, 1951; Fonagy, 1966; Lehiste, 1970) should have resulted in the accurate perception of stressed words compared to the unstressed-word condition.

4.1.2. Perception of speech rhythm

The first and second objectives of the study were to measure the rhythm perception scores of children with normal hearing (CWNH) and children with hearing impairment (CWHI) and compare between the groups. The results of the current study indicated that the perception of speech rhythm was accurate among all the age groups of the CWNH group except for one. Three of the age groups of the CWNH group were able to produce and maintain the speech rhythm as syllable-timed similar to the stimuli sentences. The age group, 4.0-4.11, was found to have produced an average nPVI of 59.9, which is closer to the lower boundary of the high nPVI classification. Also, the potential of children having a different rhythm pattern during their developmental years of language acquisition could be a typical developmental pattern that requires more evidence to be established. The perception of speech rhythm was found to be deviant in the CWHI group. The perceptions of three out of four groups of CWHI groups were not able to perceive the rhythm pattern of the stimuli accurately, and consequently, they could have produced a stress-timed rhythm pattern. The oldest age group of 6.0-6.11 years produced rhythm patterns in syllable-timed language similar to the stimuli. The results indicated that the speech stimuli presented were syllable-timed language based on nPVI and rPVI values of 43.98 and 54.72, respectively. These results were in agreement with the PVI values (syllable-timed) of the Malayalam language proposed by Savithri et al. (2006).

4.1.3. Perception of vocal emotion

The results of the study revealed that happy emotion perception was not significantly different between CWNH and CWHI, indicating that both the groups performed at similar levels. This result is in contradiction with the studies reporting the lower performance of children as an auditory deficit in comparison to their typical peer group (Most & Peled, 2007; Most & Aviner, 2009; Most, Bachar, & Dromi 2012; Most & Michaelis, 2012). The authors of the current study explain this due to the training both the groups received before the testing phase. The children with hearing impairment could perceive the minute differences in the frequency and duration attributes to accurately perceive the happy emotion even with a short duration of training. Also, the acoustic parameters (Table 2.3) of happy emotion had a higher pitch variation, wider pitch range, and greater intensity, as shown in the study by Luo, Fu, and Galvin III (2007), which could have facilitated the accurate perception of the emotion by children with hearing impairment.

While the training phase helped children with hearing impairment to achieve perception scores similar to their language age-matched typical children in the happy emotion condition, the results of the sad and neutral emotional prosody indicated a significant difference between the perception scores of CWNH and CWHI. Luo et al. (2007) pointed out that the acoustic properties of the sad emotion include a lower pitch variation, lesser pitch range, and lower intensity. Also, Yildirim et al. (2004) revealed in their study that sad and neutral emotion conditions shared similar acoustical properties. The results of the current study demonstrated that the acoustic properties of the sad and neutral emotional conditions were difficult to be perceived by children with hearing impairment even after a short-term training. Even though overt significant perception score differences were not observed, the authors report covert changes indicating improvement in the perception of the emotions along the training phase continuum. It can also be pointed out that the perception scores for all the three conditions (happy, sad, and neutral) improved from the initial to the final sessions, indicating that with training, the children with hearing impairment using hearing aids can improve their supra segmental perception abilities.

As mentioned earlier, the perception of suprasegmental features of speech like emotional prosody has a significant role in enhancing the social and linguistic development of the child (Grossman et al. 2005; Chatterjee et al. 2015). The deprival of auditory stimuli leads to a lack of comprehension of the speaker's intentions even when the content of the communication is conveyed (Mellon, 2000). This can have an impact on the linguistic and social development of the child in terms of pragmatics of speech, social responsiveness, and even cognitive qualities like sympathy, empathy, and apathy. Even the theory of mind is considered closely related to emotion perception in children (Cutting & Dunn, 1999; Laugen et al., 2016). Peterson (2009), in his study, pointed out that children with hearing impairment are at high risk for delayed development of the theory of mind due to the deficit in perception of emotions.

The study done by Kalathottukaren et al. (2017) revealed that the prosody perception and production in people with a hearing impairment has always received less attention both in assessment and intervention when compared to other aspects of speech. The current study results also point out that even with befitting hearing aids, with aided hearing thresholds within the speech spectrum, children with hearing loss pay less attention to the acoustic characteristics that have prosodic cues. Moreover, with training, children with hearing can perceive these acoustic features (Murray & Arnott, 1993; Scherer, 2003), which improves the perception of prosody.

4.2. Production of speech prosody in children with hearing impairment

4.2.1 Production of speech stress

The acoustic parameters of frequency, intensity, and duration of the primary stressed word were assessed in the sentences produced by the participants of both CWHI and CWNH groups. It was found that the intonation units produced by the participants of the CWHI group were more in number when compared to the CWNH participants. This suggests that CWHI participants produced more pauses in their speech when compared to their language-age matched peer group. Similar observations were found by Patil, Sindhura, and Reddy (2010) among a group of children with cochlear implants. They reported that children with hearing impairment inserted frequent and inappropriate, long pauses in a sentence that resulted in frequent F0 changes and hence more number of intonation units.

The fourth objective of the study was to compare the stress productions of CWNH and CWHI. There was no significant difference that was found between the fundamental frequency (F0) measures of the stressed word produced by both groups across age groups. This could be either because the CWHI participants can produce the F0 cues for stress similar to their language age-matched peer group or because F0 does not account for the stress cues in the Malayalam language. The latter should be more relevant as the studies on stress acoustics in the Malayalam language have not accounted for pitch as a dominant cue for stress (Irfana et al., 2014).

The mean intensity measures of CWHI were slightly increased compared to the mean scores of CWNH (Table 2.3). It was also observed during the training phase that the majority of the CWHI participants increased their loudness to express the stress in their speech. This supports the evidence that CWHI uses amplitude as the major acoustic cue to identify stress variations (Rubin-Spitz & McGarr, 1986; Most, 2015). There was a developmental trend that was observed for the acoustic cue of intensity in the results. The intensity measures of the CWNH and CWHI were found to be significantly different between the groups CWNH and CWHI across the age groups of 3.1-3.11; 4.0-4.11; 5.0-5.11 years. This finding could also indicate that intensity is an important acoustic cue for stress in the Malayalam language. In the age group of 6.0-6.11 years, there was no significant difference in performance found between the groups CWNH and CWHI. This suggests that the older age group of CWHI produced the stress cue of intensity similar to that of the CWNH group. This suggests that with the increase in language age, the CWHI participants were able to produce the intensity cues for stress more accurately. Further investigations aimed at finding the relation of chronological or language age with stress productions are required to infer the validity of the above observation.

The mean values of duration measures for both groups suggested that the CWHI had greater duration values than the CWNH group. This observation of CWHI prolonging their utterances, especially stressed words are in consensus with numerous studies done in different languages (Boothroyd et al. 1974; Stark & Levitt, 1974; Nickerson, 1974). The duration cues for stress were found to be significantly different for the age groups 3.1-3.11 and 4.0-4.11 years, suggesting that the CWHI participants were not able to produce the duration cues to represent the speech stress. However, no significant difference was found between the duration measures of the age groups of 5.0-5.11 and 6.0-6.11 years of CWNH and CWHI groups. This indicates that the older groups had reduced the duration measures of stressed words compared to the younger groups, and their productions were similar to that of the language age-matched CWNH peer group. Also, this may be an indication that duration is a major acoustic cue in the expression of stress in the Malayalam language. This finding is supported by the study done by Irfana et al. (2014) quoted duration as the major acoustic cue for emphatic stress in the Malayalam language. The developmental trend of the duration parameter in the stressed word condition that was obtained is presented in Figure 4.1. It was observed that the greater duration in CWHI reduced as the age increased.

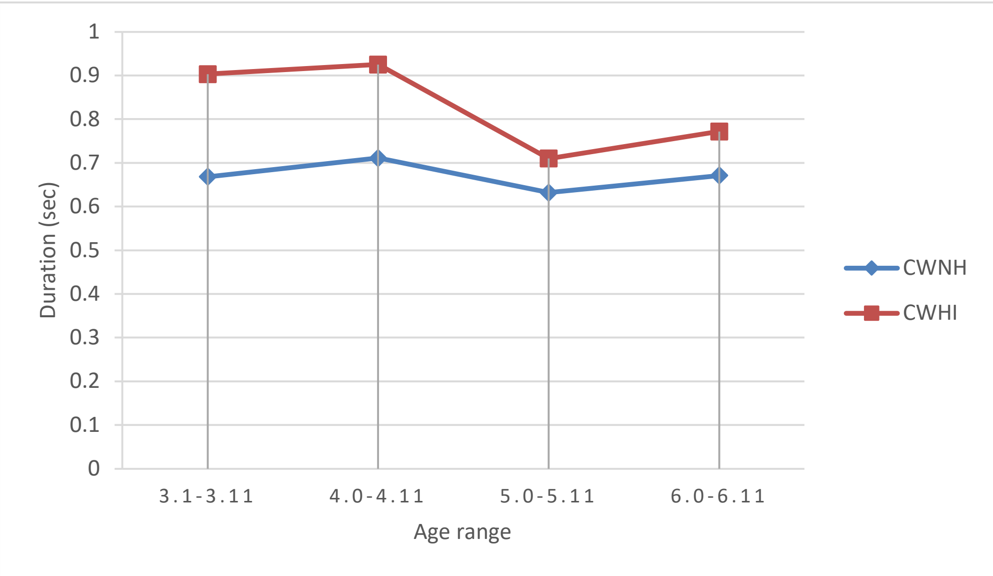

Figure 4.1 Developmental trend observed in duration parameter in the stressed-word condition in the CWNH and CWHI groups.

4.2.2. Production speech rhythm

A PVI value above 50 was considered high, and below 50 was considered low (Merin, 2016). The rhythm class of all the age groups of CWNH was syllable-timed, except for the age group of 4.0-4.11, which did not fall under any rhythm class. However, it should be noted that the mean nPVI value of this age group was 59.99, which is to the lower boundary of the nPVI-high classification. The rhythm class of all the age groups under CWHI fell under stress-timed rhythm class, except for the 6.0-6.11 age group with PVI values of syllable-timed rhythm class. The CWNH group participants were able to produce syllable-timed rhythm patterns with low nPVI and high rPVI values, similar to the rhythm pattern of native Malayalam speaking adults (Savithri et al., 2006). This indicates that CWNH participants were able to equate their speech to the adult pattern of speaking from an early age of 3 years. The rPVI values (100.23) were slightly increased, indicating the production of the consonants in their speech was stressed. Also, a developmental trend of increase in rPVI values and a decrease in nPVI values were found in the CWNH group.

The CWHI group participants (3.0-5.11 years) were found to produce stress-timed rhythm patterns with high nPVI and high rPVI values. This indicates that the CWHI participants were giving more stress to consonants and vowels, prolonging the productions were compared to CWNH group. The speech of the children with hearing impairment is reported to have inconsistent prolonged vowels (Calvert, 1962; Parkhurst & Levitt, 1978) and stressed syllables (Rosenhouse,1986). Similar observations of increased nPVI and rPVI compared to the typically developing peer group were found by Savithri et al. (2008) and Merin (2016). Also, considering the change of rhythm pattern in the 6.0-6.11 age group to syllable-timed rhythm pattern indicates that with greater duration of the training, the CWHI participants could achieve the adult-like rhythm pattern of the Malayalam language. Hidalgo, Falk, and Schon (2017) reported in their study that intensive training facilitated with the increase in the mental age as children with hearing impairment could result in better rhythm perception and production. Ramus (2002) reported that the syllabic combinations of stress-timed languages are more complex such as CCCVCCC. Then it increases in a syllable-timed language such as CCVCC, and then it is the least in mora-timed languages as in CVCV. The current study results found that the speech of the CWHI group was predominantly stress-timed, except for the eldest age group. This is contradictory to the study done by Ramus (2002), as it was observed that CWHI (3.1-6.0 years) were simplifying the syllabic structure of the stimuli in their productions and yet were producing stress-timed rhythm patterns. A similar finding of simplification of the syllable structures was found in their study by Savithri et al. (2008).

The fifth objective of the study was to formulate a developmental profile of prosody in CWNH and CWHI. Savithri et al. (2006) found that the adult native speakers of the Malayalam language have low nPVI and high rPVI values. In the current study, when the PVI values of different age groups were compared to the adult native Malayalam rhythm pattern, we can see a developmental trend of increased nPVI values and a decrease in the rPVI values evident in the CWNH group. While, in the CWHI group, we observed a developmental increase and then a decrease in both nPVI and rPVI values, moving the pattern closer to the adult PVI values. A scatter plot representation of the same is given in Figures 4.2 & 4.3.

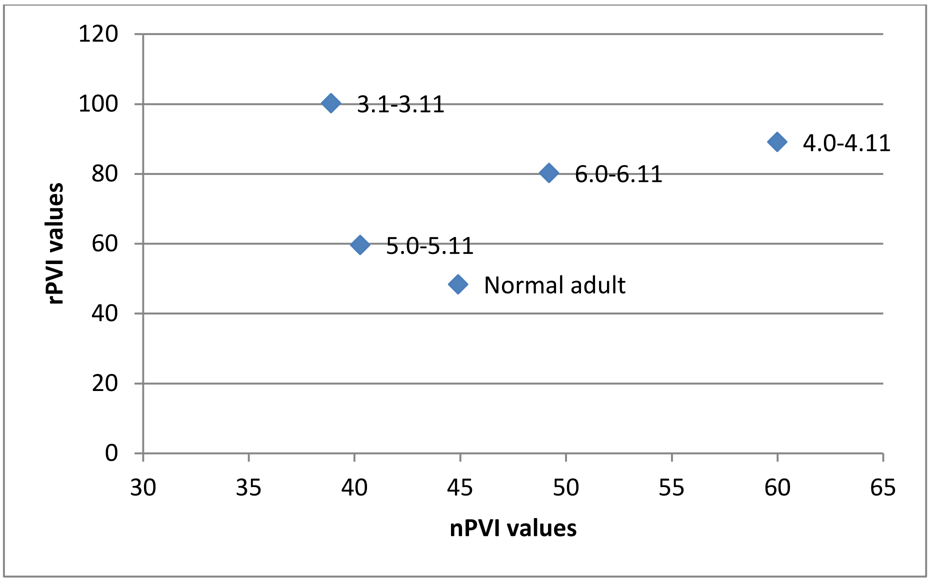

Figure 4.2. The PVI values of participants across the age groups (3.1-6.11 y) of the CWNH group.

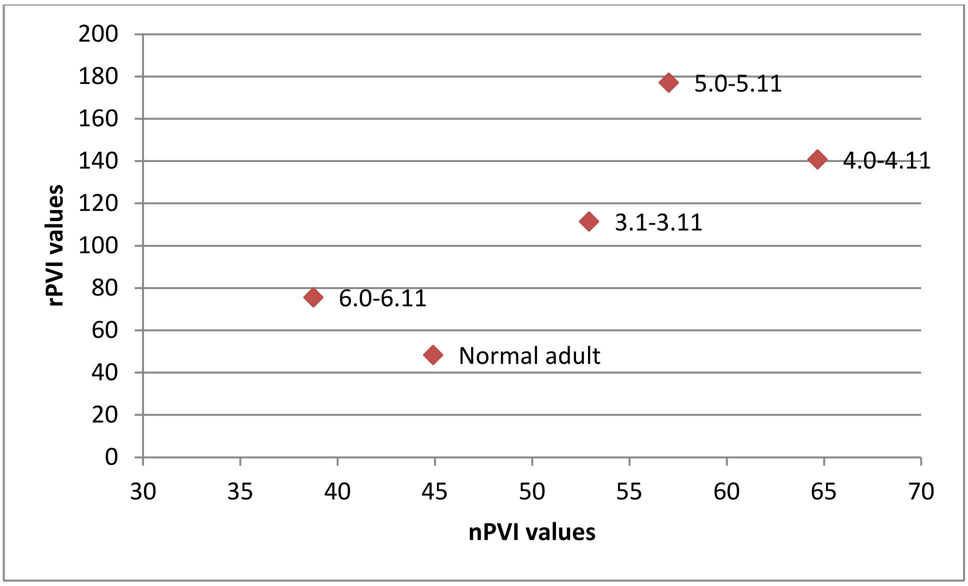

Figure 4.3. The PVI values of participants across the age groups (3.1-6.11 y) of the CWHI group.

There have been studies suggesting that children with CIs or HAs can discriminate between different rhythmic patterns (Innes-Brown, 2013) and yet the inclusion of assessment and management of these parameters are still less sought. Using training strategies for rhythm in the intervention of children with hearing impairment can help them to improve their temporal productions, which can facilitate better speech perception and production (Hidalgo, Falk & Schon, 2017). This can eventually result in a more natural speech with increased speech intelligibility.

4.2.3. Production of speech intonation

During the training and testing phases, it was observed that there was a significant difference between CWNH and CWHI in the production of happy and sad emotions. The neutral emotion was produced similarly by both the groups, maintaining the flat pitch contour.

a) Happy Emotion

The mean onset and offset frequencies of happy intonation were similar in CWNH and CWHI. There was no particular developmental trend that was observed across the age group for the onset frequency measurements. It was observed that CWHI exhibited a reduced standard deviation in comparison with the CWNH peer group. This observation could suggest that the pitch production range of CWHI is restricted compared to their typical peer group. Voelker (1935) and Greene (1965) also reported in their study that the frequency range production of deaf children was much lesser than their normal-hearing peer group.

For the onset frequency, there was a significant difference between the groups shown by the age group of 4.0-4.11 years, while there was no significant difference at the next age level, which could be an indicator that by 5.0-5.11 years, the CWHI participants were able to achieve similar onset frequency productions as their typical peer group. A significant difference was observed between the age groups 4.0-4.11 and 5.0-5.11 in the offset frequency. This could be because the CWNH and CWHI have similar productions at a young age (3.0- 3.11 years), but as the age increases (>3.11 years), CWNH may attain the ability to imitate the adult offset frequency pattern. At the same time, the CWHI fails to attain this normalcy due to the auditory deficit. As the age increased, the CWHI group was able to produce offset frequency (rising intonation) similar to their typical peer group, which could have resulted in no significant difference between the two groups (Table 3.19).

The terminal pitch contour analysis also showed that as age increased, there was an increase in the accurate production of rising intonation contour in both groups. Though CWNH achieved accurate production of rising intonation contour by 4.0-4.1 years of age, even at age 6.0-6.11 years, only 40%of CWHI groups could produce a rising intonation contour. Similar observations were reported by Philips et al. (1968), where the author found that deaf children found it difficult to produce the terminal rise in intonation in types of sentences like happy or interrogative sentences till late age. Among the CWHI participants of the current study, it was observed that they substituted the rising tone with a flat intonation contour in response. The developmental trend of the rising intonation contour is given in Figure 4.4, where it could be observed that CWHI had consistent poor production of rising contour compared to CWNH across age groups.

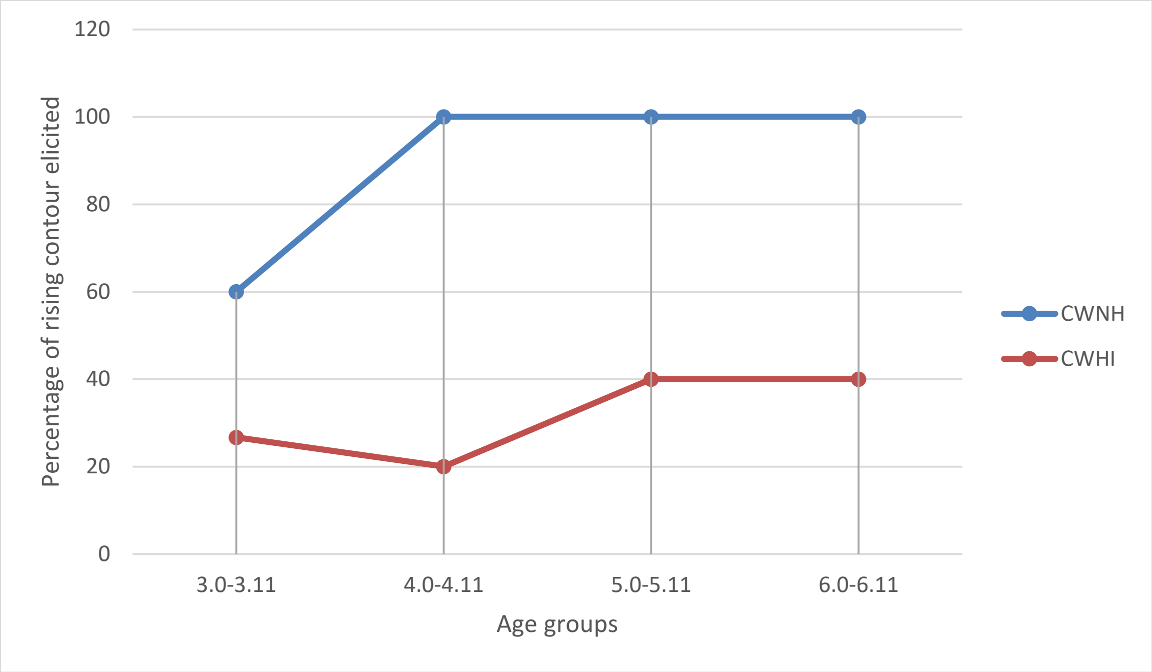

Figure 4.4 The developmental trend of rising intonation contour for happy emotion by CWNH and CWHI groups

b) Sad Emotion

The results of descriptive statistics showed that the mean onset frequency values were similar between both groups for the sad intonation (Table 3.20). There was a significant difference shown by the age group 4.0-4.11 years, which would be an indicator of the group differences at this age range. The observation of no significant difference at the next age level could indicate that by 5.0-5.11 years, the CWHI participants were able to achieve similar onset frequency productions as their typical peer group. The offset frequency mean values showed a significant difference in the age groups of 4.0-4.11 and 5.0-5.11. The developmental trend is observed here in Figure 4.5. It was observed that with an increase in age, there was better production of falling contour in both groups.

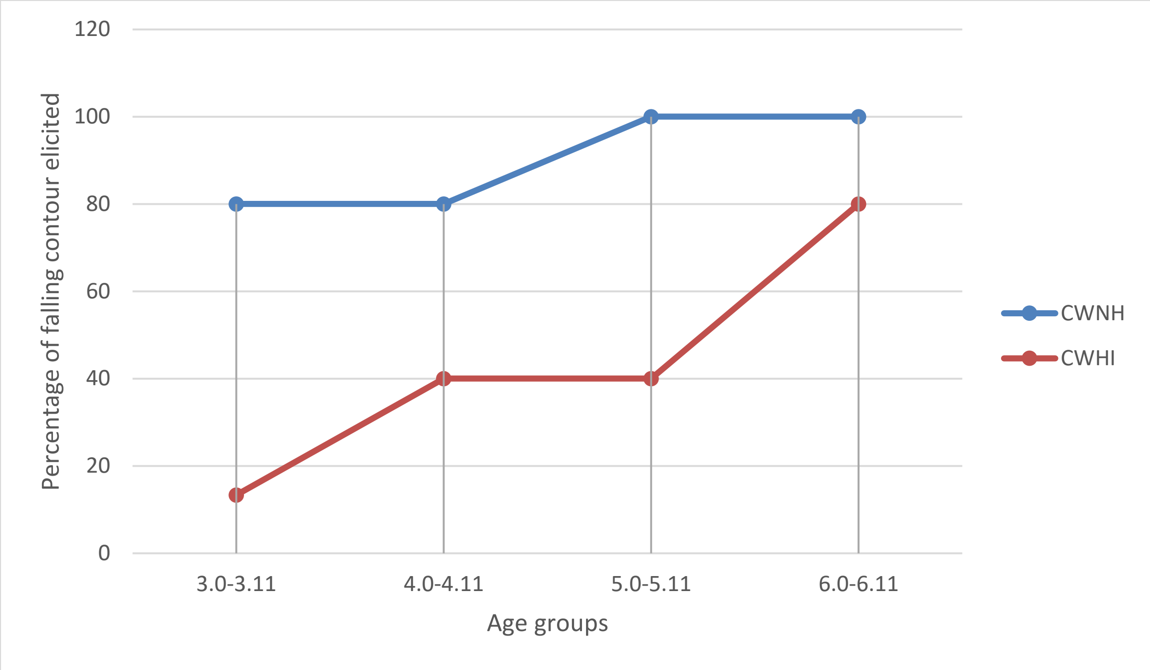

Figure 4.5 The developmental trend of falling intonation contour for sad emotion by CWNH and CWHI groups

In the case of correct terminal contour production, compared to the production of rising contour, the CWHI participants had a better percentage of correct falling intonation contour productions (Table 3.24). Many studies have reported that children achieve falling intonation contour earlier than the rising one (Philips et al., 1968). In a study with four-year-old normal-hearing children, it was found that children used more falling intonation than the adult modeled rising contours (Snow, 1998). Studies on the intonation contours of children with hearing aids also have shown that they could produce falling intonation easily, and it was easier than the rising one (Abberton et al., 1991; Most & Frank, 1994).

c) Neutral Emotion

Both groups performed similarly on the production of neutral emotion contour. It was observed that the mean onset, offset frequency values, and the intonation contour was similar in both groups (Table 3.25, 3.27, 3.29). There was no significant difference found between the groups on the mean onset frequency values. At the same time, there was a significant difference observed for the age group of 3.0-3.11 years between the groups on mean offset frequency (Table 3.27). This difference could be because of the falling intonation contours produced by 3.0-3.11-year-old CWNH participants for the neutral emotion. The CWHI participants had no difficulty maintaining their pitch level to a flat pitch contour, while young CWNH (3.0-3.11 years) had difficulty (Table 3.29).

In general, the production of vocal emotional prosody showed that the CWHI group had difficulty producing the variations or the required number of variations in their pitch contour compared to their typically developing language-aged peer group. It was observed that in emotional contexts where the CWHI were not able to produce the adequate amount of pitch variations, it either fell above the required rise in pitch or below the optimum falling contour. Similar observations were reported by Allen and Arndorfer (2000), where they found that the normal hearing talkers' contours fell further, and their rising contours rose further in comparison to the talkers with hearing impairment. Though the hearing aids can process the intonation cues and make them available for the users (Most & Peled, 2007), the performance of CWHI on intonation perception and production tasks is reported to be poorer. During the training phase of the current study, it was observed that substantial covert changes in both perception and production of both rising and falling pitch variations towards the final sessions of the training phase, indicating that the training was beneficial. There is a need to focus on the perception and production of intonation during assessment and management of the CWHI group equal to the importance given to segmental learning.

When studies have reported that hearing aids enable perception of low-frequency spectral information (Henry & Turner, 2003; Kong, Stickney & Zeng, 2005), the children with hearing impairment with befitting hearing aids show poorer performance on perception and production of speech prosody. But during the training phase of the current study, there was an improvement in both the perception and production performance of the CWHI participants observed. Though overt statistical changes were not observed, covert improvements in CWHI participants' responses were noticed, which could be the effect of the short-term prosody training that they received. Further investigations are required to evaluate the effect of prosody training on CWHI performance on prosody tasks. The covert improvement of CWHI shows that with intensive training, the perception and production of prosody can be improved in children wearing befitting hearing aids. Kalathottukaren et al. (2017) reported that clinical assessment and speech, language, and aural rehabilitation should expand to address the prosodic difficulties in perception and production in children with hearing impairment. Aural rehabilitation has always given importance to training segmental features of speech, neglecting the importance of supra-segmental features in daily life communication. There is an increasing need to give equal focus on both segmental and supra-segmental aspects of speech to improve the speech intelligibility and, subsequently, the naturalness of speech of children with hearing impairment. The developmental trend of the production of neutral intonation contour is presented in Figure 4.6. The monotone of CWHI helped them maintain the neutral intonation contour in all age groups, while only elder CWNH maintained the flat contour.

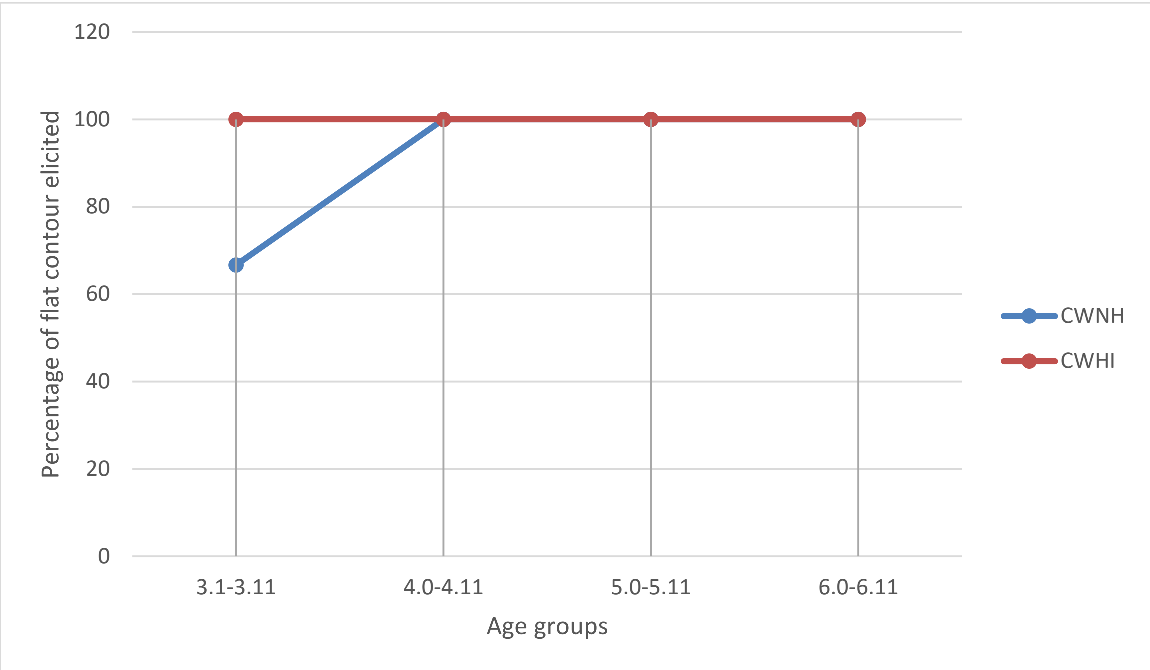

Figure 4.6 The developmental trend of flat intonation contour for neutral emotion by CWNH and CWHI groups

CHAPTER 5

SUMMARY AND CONCLUSION

The aim of the study was to investigate the perception and production of prosody (stress, rhythm, and intonation) in children with hearing impairment. The different objectives taken up were:

1. To measure the perception of stress, intonation, and rhythm in children with hearing impairment and their language-age matched typically developing children.
2. To compare the perception of stress, intonation, and rhythm in children with hearing impairment and their language-age matched typically developing children.
3. To measure the production of stress, intonation, and rhythm among children with hearing impairment and their language-age matched typically developing children.
4. To compare the production of stress, intonation, and rhythm among children with hearing impairment and their language-age matched typically developing children.
5. To formulate a developmental profile for prosody in children with hearing impairment.

To study the above-mentioned objectives, 30 children(16 males and 14 females) with normal hearing sensitivity (CWNH) and 30 children(18 males and 11 females) with hearing impairment (CWHI) with a language age of 3.1-6.11 years was recruited for the study. Twenty-four commonly occurring sentences from the kindergarten and lower primary school Malayalam textbooks were shortlisted by the authors. The selected sentences were concrete and representable in pictures with a sentence length of 3 words. Picture cards (one pair-stressed and unstressed per sentence) were custom drawn for each sentence, representing the unstressed/stressed words in the sentence. The 24 sentences with the two variations (stressed and unstressed) and three different emotional variations - happy, sad, and neutral; were recorded in a sound-treated room by a female speaker (the research officer) using a unidirectional microphone kept at six inches distance. To reduce the task load on the paediatric participants with regard to attention to the activities carried out, the stressed-word condition of a sentence overlapped with the happy intonation and the unstressed-word condition overlapped with the neutral intonation.

The study was carried out in two phases for both the control group and the experimental group. The demographic detail of the participants, along with the consent for participation was taken from the caretakers. The participants were allotted 180 minutes of prosody perception training before evaluating their perception of the stimuli sentences. Phase 1 of the study included the training followed by testing in phase 2.

Phase 1: The training phase was carried out in a sound-treated audiometric room with a single room set. The auditory stimuli (14 recorded sentences categorized for training) were presented through calibrated loudspeakers at 50 dB HL at a distance of 1 meter and the azimuth of 0 degrees, routed through an audiometer along with the picture representation of each sentence. The examiner was seated next to the participant, thus allowing the examiner to control the presentation of the auditory stimulus, picture cards, and smiley placards.

Phase 2: The testing was carried out in the same setup as in phase 1 as a single sitting or two sittings (according to the co-operation of the participant). The testing duration varied from 30minutes to one hour, according to the participant's response. During phase 2, the set of 10 recorded stimuli sentences (with the stress and intonation variations- 30sentences) were presented in randomized order. A maximum of two trials was given to the child if required. A forced-choice response was followed. Each child was asked to specify their response by pointing to the appropriate picture and to select the smiley placard after listening to the auditory stimuli. The correct responses were given a score of 1, and wrong responses were given a score of 0. The response of each child was tabulated for the ten stimuli sentences for each emotion.

The perception responses were tabulated for each parameter: stress, rhythm, and intonation. The responses for production tasks were fed into Praat software version 5.2.01 developed by Boersma and Weenink (2009) and were analyzed for the different parameters under stress, rhythm, and intonation. Both perception and production scores/values were subjected to suitable statistical analyses using SPSS version 20. A normality test using Shapiro Wilk's test was performed for all data sets, which determined the use of parametric or non-parametric tests. Independent t-test (parametric) or Mann-Whitney U test (non-parametric) was used for comparisons between CWNH and CWHI, while paired t-test (parametric) or Wilcoxon signed-rank test was used for pre-training and post-training comparisons within groups.

The results of the study, when discussed according to the objectives, revealed that:

1. ***To measure the perception of stress, rhythm, and intonation in children with normal hearing (CWNH) and children with hearing impairment (CWHI).***

The perception of speech stress was observed to be better in stressed-word conditions than in unstressed-word conditions for CWNH and CWHI groups. Comparing the means of both the groups, it was observed that the perception of CWHI is poorer compared to CWNH, especially in unstressed-word conditions. The rhythm perception of CWNH enabled them to produce rhythm patterns similar to that of adults (syllable-stress pattern), except for the age range of 4.0-4.11 years; while that of the CWHI had difficulty in replicating the rhythm pattern of the stimuli, as a consequence of the deficient auditory feedback. The perception of the three emotional variations: happy, sad, and neutral, was assessed for intonation perception. It was found that both CWNH and CWHI had better perception scores for happy emotion, while the scores were lowest for neutral emotion when compared within the groups. The mean perception scores revealed that CWNH and CWHI performed similarly in happy emotion identification. In contrast, the perception scores of CWHI were lower in sad and neutral emotion conditions in comparison with the CWNH group. The acoustic properties of the happy emotion would have enabled an easier identification of this emotion in comparison with sad and neutral variations.

2. To compare the perception of stress, intonation, and rhythm in children with hearing impairment and their language-age matched typically developing children.

The CWNH participants were able to perform better on the stress perception task than CWHI. Both the groups had lower scores on the unstress-word condition compared to the stress-word condition. There was a significant difference found between the perception scores of stress and unstress of both the groups. The intonation perception of happy emotion was found to have no significant difference, while sad and neutral emotion showed a significant difference. The rhythm perception of CWNH and CWHI was found to be different, as the CWNH participants were able to replicate the syllable stress pattern of the stimuli.

3. To measure the production of stress, intonation, and rhythm among children with hearing impairment and their language-age matched typically developing children.

In the production of speech stress, the frequency and intensity mean values were observed to be similar across the age groups. While the duration mean values were observed to be greater for CWHI compared with CWNH, the stressed words were prolonged in the CWHI group. The rhythm measurements showed that the PVI values of CWNH match that of the adult rhythm pattern. The CWHI PVI values were greater than CWNH and were categorized under the unclassified rhythm class.

In the production of intonation patterns, the onset frequency mean values were similar between CWNH and CWHI for all the emotional variations. The offset frequency was significantly different in the two middle-age groups (4.0-4.11years and 5.0-5.11 years) for happy and sad emotional variations. The CWNH participants were able to produce the terminal contour accurately in all the emotional variations. In contrast, the participants of CWHI had difficulty maintaining the terminal contours of happy and sad emotions.

4. To compare the production of stress, intonation, and rhythm among children with hearing impairment and their language-age matched typically developing children.

Speech stress production tasks showed no significant difference between the fundamental frequency (F0) measures of the stressed word produced by both groups across age groups. In comparison, there was a significant difference that was found between the intensity and duration measurements. The speech intonation production tasks showed that for the happy emotion, there was no significant difference between the CWNH and CWHI groups for the onset and offset frequency. At the same time, the CWHI participants had difficulty producing the rising intonation contour compared to the CWNH group. The flat intonation contour was produced accurately by both groups. The offset frequency showed a significant difference between the groups for sad and neutral emotion production, while onset frequency showed no significant difference between the groups. The rhythm productions were found to be different for both CWNH and CWHI across the age groups. The rhythm pattern produced by CWNH was categorized as syllable-timed rhythm pattern, while that of the CWHI was found to be unclassified rhythm class, except for the elder age group of 6.0-6.11 years.

5. To formulate a developmental profile for prosody in children with hearing impairment.

Speech Stress: There was an increase in the perception of stress as the language age increased in the CWHI group. In the case of the production of speech stress, it was found that there was no developmental trend seen in the production of stress for the frequency parameter. The intensity parameter observed that by the age of 6.0 years, the CWHI participants performed similarly to the CWNH participants. And for the parameter of duration, by the age of 5 years, the CWHI duration productions were similar to the CWNH productions.

Speech Rhythm: There was a developmental trend observed in the production of speech rhythm. It was found that CWHI participants in the age range of 3.0-5.11 years produced a different type of speech rhythm, which was unclassified. However, the older age group, 6.0-6.11years participants, produced the adult-like speech rhythm pattern, the syllable-timed rhythm pattern.

Speech Intonation: There was an increase in the intonation perception scores with an increase in the language age of the CWHI participants. There was no developmental trend in producing happy emotion parameters: onset frequency, offset frequency, and duration. The onset and offset frequencies of sad and neutral emotions showed a mixed trend. For the onset and offset frequencies of sad and neutral emotions, there was a significant difference at 4.0-4.11 years, while there was no significant difference at the next age level, which could be an indicator that by 5.0-5.11 years, the CWHI participants were able to achieve similar onset frequency productions as their typical peer group. A significant difference was observed between the age groups 4.0-4.11 and 5.0-5.11 in the offset frequency. This could be because the CWNH and CWHI had similar productions at a young age (3.0- 3.11 years), but as the age increased (>3.11 years), CWNH may attain the ability to imitate the adult offset frequency pattern. At the same time, the CWHI fails to attain this normalcy due to the auditory deficit. As the age increased, the CWHI group was able to produce offset frequency (rising intonation) similar to their typical peer group, which could have resulted in no significant difference between the two groups.

5.1 Clinical Implications

When studies have reported that hearing aids enable perception of low-frequency spectral information (Henry & Turner, 2003; Kong, Stickney & Zeng, 2005), the children with hearing impairment with befitting hearing aids show poorer performance on perception and production of speech prosody. During the training phase of the current study, it was observed that there was an improvement in both the perception and production performance of the CWHI participants. Though overt statistical changes were not observed, the investigators report covert improvements in CWHI participants' responses, which could be the effect of the short-term prosody training that they received. Further investigations are required to evaluate the effect of prosody training on CWHI performance on prosody tasks. Aural rehabilitation has represented segmental features of speech, ignoring the importance of supra-segmental features in everyday speech communication. In conclusion, there is an increasing need to give equal focus on both segmental and supra-segmental aspects of speech to improve speech intelligibility and the naturalness of speech of children with hearing impairment. Thus, it can be put forth that conventional auditory training and speech-language interventions that focus majorly on the segmental aspects of speech should expand and include training children with hearing loss on perception and production of supra segmental features of speech as well.

5.2 Limitations of the study

The current study, however, had certain limitations. A psychological screening test could have been administered to the study participants to rule out any psychological factors that may have contributed to the test performance. Also, the participants' socio-economic status was not considered during the selection, which could have been a variable in the performance on the prosody perception and production tasks. Subjective measures were used to perceive speech prosody, while objective measures have been derived from producing speech prosody. Also, a specific rhythm perception training was not materialized by the examiners. It was assumed that the participants who were able to replicate the rhythm pattern of the stimuli were considered to have perceived the rhythm pattern of the stimuli correctly. The age range selected for the study was constricted due to the unavailability of participants. Likewise, to bring better clarity to the developmental trend observed in the current study, more participants should have been recruited in each age range

REFERENCES

Abberton, E., Fourcin A. and Hazan, V. (1991).Fundamental frequency range and the development of intonation in a group of profoundly deaf children.*Proceedings of the XIIth International Congress of Phonetic Sciences, Aix-en-provence*, Volume 5.

Abercrombie, D. (1964). *Syllable quantity and enclitics in English*.na.

Abercrombie, D. (1967). *Elements of general phonetics* (Vol. 203). Edinburgh: Edinburgh University Press.

Alkhamra, R. A., & Abu-Dahab, S. M. (2020). Sensory processing disorders in children with hearing impairment: Implications for multidisciplinary approach and early intervention. *International Journal of Pediatric Otorhinolaryngology*, *136*, 110154.

Allen, G. D., & Arndorfer, P. M. (2000). Production of sentence-final intonation contours by hearing-impaired children. *Journal of Speech, Language, and Hearing Research*, *43*(2), 441-455.

Anderson F. (1960), An experimental pitch indicator for training deaf scholars, Jornal of American Society of Acoustics, 32(8), 1065-1074.

Angelocci, A. A. (1962). Some observations on the speech of the deaf. *Volta Review*, *64*(7), 403-405.

Arciuli, J., Monaghan, P., & Seva, N. (2010).Learning to assign lexical stress during reading aloud: Corpus, behavioral, and computational investigations. *Journal of Memory and Language*, *63*(2), 180-196. https://doi.org/10.1016/j.jml.2010.03.005

Armony, J. L., Chochol, C., Fecteau, S., & Belin, P. (2007). Laugh (or cry) and you will be remembered: influence of emotional expression on memory for vocalizations. *Psychological Science*, *18*(12), 1027-1029.

Arvaniti, A. (2009). Rhythm, timing and the timing of rhythm. *Phonetica*, *66*(1-2), 46-63.

Asher, R. E. (2013). *Malayalam*.Routledge.https://doi.org/10.4324/9781315002217

Aslin, R. N., & Newport, E. L. (2008). What statistical learning can and can't tell us about language acquisition. In *Infant pathways to language* (pp. 33-48).Psychology Press.

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*. https://doi.org/10.1044/1092-4388(2012/11-0257)

Banse, R., & Scherer, K. R. (1996).Acoustic profiles in vocal emotion expression. *Journal of personality and social psychology*, *70*(3), 614.

Barker, D. H., Quittner, A. L., Fink, N. E., Eisenberg, L. S., Tobey, E. A., Niparko, J. K., & CDaCI Investigative Team (2009). Predicting behavior problems in deaf and hearing children: the influences of language, attention, and parent-child communication. *Development and psychopathology*, *21*(2), 373–392. https://doi.org/10.1017/S0954579409000212

Barzaghi, L., & Mendes, B. (2008, May). Stressed and unstressed vowel production in hearing-impaired speech. In *Speech prosody conference: Proceedings. Brazil: Campinas* (pp. 109-202).

Besson, M., Magne, C., & Schön, D. (2002). Emotional prosody: sex differences in sensitivity to speech melody. *Trends in cognitive sciences*, *6*(10), 405-407. https://doi.org/10.1016/S1364-6613(02)01975-7

Blamey PJ, Sarant JZ, Paatsch LE, Barry JG, Bow CP, Wales RJ, Wright M, Psarros C, Rattigan K, & Bloomfield, L. (1933). *Language history: from Language (1933 ed.)*. Holt, Rinehart & Winston.

Boersma, P., & Weenink, D. (2009).Praat: doing phonetics by computer. *Computer program available at: http://www. praat. org*.

Bolinger, D. L. (1958).Stress and information. *American Speech*, *33*(1), 5-20. https://doi.org/10.2307/453459

Bolton, T. L. (1894). Rhythm. *The american journal of psychology*, *6*(2), 145-238.

Boothroyd, A. (1984). Auditory perception of speech contrasts by subjects with sensorineural hearing loss. *Journal of Speech, Language, and Hearing Research*, *27*(1), 134-144. https://doi.org/10.1044/jshr.2701.134

Boothroyd, A., & Eran, O. (1994). Auditory speech perception capacity of child implant users expressed as equivalent hearing loss. *Volta Review*, *96*(5), 151-67.

Bradley, M. M., Greenwald, M. K., Petry, M. C., & Lang, P. J. (1992).Remembering pictures: pleasure and arousal in memory. *Journal of experimental psychology: Learning, Memory, and Cognition*, *18*(2), 379.

Calvert, D. (1962). Deaf voice quality: a preliminary investigation. Volta Review, 64: 402-403

Carroll J, Zeng FG. (2007) Fundamental frequency discrimination and speech perception in noise in cochlear implant simulations. Hear Res 231(1–2):42–53.

Chatterjee, M., Zion, D. J., Deroche, M. L., Burianek, B. A., Limb, C. J., Goren, A. P., ... & Christensen, J. A. (2015). Voice emotion recognition by cochlear-implanted children and their normally-hearing peers. *Hearing research*, *322*, 151-162.https://doi.org/10.1016/j.heares.2014.10.003

Clement, C. J. (2004). *Development of vocalizations in deaf and normally hearing infants*.Netherlands Graduate School of Linguistics.

Clement, C.J., den Os, E.A. & Koopmans-van Beinum, F.J. (1996).The development of vocalizations of deaf and normally hearing infants. In T.W. Powell (Ed.) *Pathologies of Speech and Language: Contributions of Clinical Phonetics and Linguistics*. A publication of the International Clinical Phonetics and Linguistics Association.

Cohen-Licht, D., & Most, T. (2000). The contribution of speech envelope to the perception of intonation by hearing and hearing impaired individuals. *Unpublished master's thesis). Tel Aviv University, Tel Aviv, Israel.(Hebrew)*.

Cristia, A., & Seidl, A. (2011). Sensitivity to prosody at 6 months predicts vocabulary at 24 months. In *BUCLD 35: Proceedings of the 35th annual Boston University Conference* *on Language Development* (pp. 145-156). Cascadilla Press. Retrieved from: https://pure.mpg.de/rest/items/item_1094660/component/file_1105580/content

Curtin, S. (2010). Young infants encode lexical stress in newly encountered words. *Journal of experimental child psychology*, *105*(4), 376-385.https://doi.org/10.1016/j.jecp.2009.12.004

Curtin, S., Mintz, T. H., & Christiansen, M. H. (2005). Stress changes the representational landscape: Evidence from word segmentation. *Cognition*, *96*(3), 233-262. https://doi.org/10.1016/j.cognition.2004.08.005

Cutting, A. L., & Dunn, J. (1999). Theory of mind, emotion understanding, language, and family background: Individual differences and interrelations. *Child development*, *70*(4), 853-865.

Davis, B. L., MacNeilage, P. F., & Matyear, C. L. (2002). Acquisition of serial complexity in speech production: A comparison of phonetic and phonological approaches to first word production. *Phonetica*, *59*(2-3), 75-107. https://doi.org/10.1159/000066065

De Clerck, I., Verhoeven, J., Gillis, S., Pettinato, M., & Gillis, S. (2019). Listeners' perception of lexical stress in the first words of infants with cochlear implants and normally hearing infants. *Journal of communication disorders*, *80*, 52-65. https://doi.org/10.1016/j.jcomdis.2019.03.008

De Rosnay, M., Pons, F., Harris, P. L., & Morrell, J. M. (2004). A lag between understanding false belief and emotion attribution in young children: Relationships with linguistic ability and mothers' mental-state language. *British Journal of Developmental Psychology*, *22*(2), 197-218.

De Sonneville, L. M. J., Verschoor, C. A., Njiokiktjien, C., Op het Veld, V., Toorenaar, N., & Vranken, M. (2002). Facial identity and facial emotions: speed, accuracy, and processing strategies in children and adults. *Journal of Clinical and experimental neuropsychology*, *24*(2), 200-213.

Deepa, M. S., Shyamala, K. C., & Deepthi, K. J. (2014).Modified Receptive and Expressive Language Test (M-RELT) for children between three to seven years. Project funded by AIISH Research Fund. Mysuru: All India Institute of Speech and Hearing.

DePaolis, R. A., Vihman, M. M., & Kunnari, S. (2008). Prosody in production at the onset of word use: A cross-linguistic study. *Journal of Phonetics*, *36*(2), 406-422. https://doi.org/10.1016/j.wocn.2008.01.003

Dhillon, R. K. (2010). *Stress and tone in Indo-Aryan languages*. Yale University.

Duffy, J. R. (2005). Motor Speech Disorders: Substrates. *Differential Diagnosis, and Management*, (2nd ed., pp. 187-215). St Louis. MO: Elsevier Mosby.

Dupuis, K., & Pichora-Fuller, M. K. (2014).Intelligibility of emotional speech in younger and older adults. *Ear and hearing*, *35*(6), 695-707.

Dyck, M. J., Farrugia, C., Shochet, I. M., & Holmes-Brown, M. (2004). Emotion recognition/understanding ability in hearing or vision-impaired children: do sounds, sights, or words make the difference?. *Journal of Child Psychology and Psychiatry*, *45*(4), 789-800.

Echols, C. H., & Newport, E. L. (1992).The role of stress and position in determining first words. *Language acquisition*, *2*(3), 189-220.https://doi.org/10.1207/s15327817la0203_1

Erber, N. P. (1979). Speech perception by profoundly hearing-impaired children. *Journal of speech and hearing disorders*, *44*(3), 255-270.https://doi.org/10.1044/jshd.4403.255

Erber, N. P. (1972b) Speech-envelope cues as an acoustic aid to lipreading for profoundly deaf children. Journal of Acoustical Society of America, 51(4):1224–1227.

Estes, K. G., & Bowen, S. (2013). Learning about sounds contributes to learning about words: Effects of prosody and phonotactics on infant word learning. *Journal of experimental child psychology*, *114*(3), 405-417.https://doi.org/10.1016/j.jecp.2012.10.002

Flaugnacco, E., Lopez, L., Terribili, C., Zoia, S., Buda, S., Tilli, S., ...& Schön, D. (2014). Rhythm perception and production predict reading abilities in developmental dyslexia. *Frontiers in human neuroscience*, *8*, 392.

Fónagy, I. (1966). Electrophysiological and acoustic correlates of stress and stress perception. *Journal of speech and hearing research*, *9*(2), 231-244. Retrieved from: https://jshd.pubs.asha.org/doi/pdf/10.1044/jshr.0902.231?casa_token=bk0kQxXFReQAAAAA:Vmu5tdfkVoWhFPrAbGw4WQk6CC1nOxBdAixNxLebsdfeWG0lQPt-xpB_swtTxwLPpso9FNlWgZGq-KY9

Frazier, L., Carlson, K., & Clifton Jr, C. (2006). Prosodic phrasing is central to language comprehension. *Trends in cognitive sciences*, *10*(6), 244-249.https://doi.org/10.1016/j.tics.2006.04.002

Gibbon, D., & Gut, U. (2001).Measuring speech rhythm.In *Seventh European Conference on Speech Communication and Technology*.

Grabe, E., Post, B., & Watson, I. (1999, August).The acquisition of rhythmic patterns in English and French.In *Proceedings of the 14th International Congress of Phonetic Sciences* (pp. 1201-1204). Berkeley, CA: University of California.

Grabe, E., & Low, E. L. (2002). Durational variability in speech and the rhythm class hypothesis. In Gussenhoven, C., & Warner, N. (2006) Eds, *Laboratory Phonology, 7*, 515-546. Berlin: Mouton de Gruyter

Gray, C., Hosie, J., Russell, P., Scott, C., & Hunter, N. (2007).Attribution of emotions to story characters by severely and profoundly deaf children. *Journal of Developmental and Physical Disabilities*, *19*(2), 145-159.

Green, D. (1956).Fundamental frequency of the speech of profoundly deaf individuals.Unpublished Doctoral dissertation,Purdue University, West Lafayette, IN

Grossmann, T., Striano, T., &Friederici, A. D. (2005).Infants' electric brain responses to emotional prosody. *Neuroreport*, *16*(16), 1825-1828. https://doi.org/10.1097/01.wnr.0000185964.34336.b1

Hall, G. S., &Jastrow, J. (1886).Studies of rhythm. *Mind*, *11*(41), 55-62.

Harris, P. L., de Rosnay, M., & Pons, F. (2005).Language and children's understanding of mental states. *Current directions in psychological science*, *14*(2), 69-73.

Harrison, M., & Roush, J. (1996). Age of suspicion, identification, and intervention for infants and young children with hearing loss: A national study. *Ear and hearing*, *17*(1), 55-62.

Hayes, B. (1995). *Metrical stress theory: Principles and case studies*. University of Chicago Press.

Henry, B. A., & Turner, C. W. (2003).The resolution of complex spectral patterns by cochlear implant and normal hearing listeners. Journal of the Acoustical Society of America, 113, 2861–2873.

Hermans, D., Houwer, J. D., &Eelen, P. (1994). The affective priming effect: Automatic activation of evaluative information in memory. *Cognition & Emotion*, *8*(6), 515-533.

Hidalgo, C., Falk, S., &Schön, D. (2017). Speak on time! Effects of a musical rhythmic training on children with hearing loss. *Hearing research*, *351*, 11-18.

Hirsch-Pasek, K., Kemler Nelson, D. G., Jusczyk, P. W., Wright Cassidy, K., Druss, B., & Kennedy, L. (1987). Clauses are perceptual units for prelinguistic infants. Cognition, 26, 269–286.

Innes-Brown, H., Marozeau, J. P., Storey, C. M., & Blamey, P. J. (2013).Tone, rhythm, and timbre perception in school-age children using cochlear implants and hearing aids. *Journal of the American Academy of Audiology*, *24*(9), 789-806.

Irfana, M., Rofina, B., & Sreedevi, N. (2014). Acoustic correlates of emphatic stress in Malayalam. *Journal of all india institute of speech and hearing*, *59*, 99-105. Retrieved from: http://www.aiishmysore.com/en/pdf/JAIISH_Vol-33.pdf#page=64

Jackson, C. W., & Schatschneider, C. (2014).Rate of language growth in children with hearing loss in an auditory-verbal early intervention program. *American annals of the deaf*, *158*(5), 539-554.

Jackson, P., & Kelly-Ballweber, D. (1986).The relationship between word and stress pattern recognition ability and hearing level in hearing-impaired young adults. *The Volta Review, 88*(6), 279-287.

Jassem, W. (1959).The phonology of Polish stress. *Word*, *15*(2), 252-269. https://doi.org/10.1080/00437956.1959.11659698

Jaya, P. (1992). Stress: Development in Tamil speaking children. (Unpublished master‟s dissertation). All India Institute of Speech and Hearing, Mysore, Karnataka.

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.

Johnson, E. K., & Jusczyk, P. W. (2001). Word segmentation by 8-month-olds: When speech cues count more than statistics. *Journal of memory and language*, *44*(4), 548-567.

Kaipa, R., & Danser, M. L. (2016). Efficacy of auditory-verbal therapy in children with hearing impairment: A systematic review from 1993 to 2015. *International journal of pediatric otorhinolaryngology*, *86*, 124-134.

Kalathottukaren, R. T., Purdy, S. C., & Ballard, E. (2017).Prosody perception and production in children with hearing loss and age-and gender-matched controls. *Journal of the American Academy of Audiology*, *28*(4), 283-294. https://doi.org/10.3766/jaaa.16001

Keane, E. (2006). Prominence in Tamil. *Journal of the International Phonetic Association*, *36*(1), 1-20. https://doi.org/10.1017/S0025100306002337

Kensinger, E. A., & Corkin, S. (2003). Memory enhancement for emotional words: Are emotional words more vividly remembered than neutral words?. *Memory & cognition*, *31*(8), 1169-1180.

Kong YY, Stickney GS, & Zeng FG.(2005) Speech and melody recognition in binaurally combined acoustic and electric hearing. J AcoustSoc Am 117(3 Pt 1):1351–1361.

Kuhl, P. K. (2000). A new view of language acquisition. *Proceedings of the National Academy of Sciences*, *97*(22), 11850-11857.

Kumar, R. B., & Bhat, J. S. (2009). Acoustic correlates of stress in Konkani language. *Language in India*, *9*(4). Retrieved from: http://www.languageinindia.com/april2009/konkanispeech.pdf

Lacheret-Dujour, A. (2015). Prosodic clustering in speech. *Emotion in Language: Theory–research–application*, *10*, 175.

Laugen, N. J., Jacobsen, K. H., Rieffe, C., & Wichstrøm, L. (2016).Emotion understanding in preschool children with mild-to-severe hearing loss. *The Journal of Deaf Studies and Deaf Education*, *22*(2), 155-163.

Lederberg A. R, Prezbindowski A. K, & Spencer PE. (2000) Word learning skills of deaf preschoolers: the development of novel mapping and rapid word-learning strategies. Child Dev 71(6): 1571–1585.

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

Lehiste, I. (1970). Suprasegmentals. Oxford, England: Massachusetts Inst. of Technology P.
Lehiste, I., &Ivić, P. (1973).Interaction between tone and quantity in Serbocroatian. *Phonetica*, *28*(3-4), 182-190. https://doi.org/10.1159/000259455

Lieberman, P. (1960). Some acoustic correlates of word stress in American English. *The Journal of the Acoustical Society of America*, *32*(4), 451-454. https://doi.org/10.1121/1.1908095

Luo, X., Fu, Q. J., & Galvin III, J. J. (2007). Cochlear implants special issue article: Vocal emotion recognition by normal-hearing listeners and cochlear implant users. *Trends in amplification*, *11*(4), 301-315.

Maassen, B., & Povel, D. (1984).The effect of correcting fundamental frequency on the intelligibility of deaf speech and its interaction with temporal aspects.*Journal of theAcoustical Society of America*, *76*, 1673–1681.

Manjula, R. (1979). Intonation in Kannada: Some aspects. Unpublished M.Sc dissertation, University of Mysore.

Mártony, J. (1968). On the correction of the voice pitch level for severely hard of hearing subjects. *American Annals of the Deaf*, 195-202.

Massicotte-Laforge, S., & Shi, R. (2015). The role of prosody in infants' early syntactic analysis and grammatical categorization. *The Journal of the Acoustical Society of America*, *138*(4), EL441-EL446.

Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, *29*(2), 143-178.

Mellon, N. K. (2000). Psychosocial development of children in deafness. *Cochlear implants: Principles and practices*, 319-321.

Merin, J. (2016). Speech rhythm in Malayalam-speaking children with hearing impairment. (Unpublished master‟s dissertation). All India Institute of Speech and Hearing, Mysore, Karnataka.

Metz, D., Schiavetti, N., Samar, V., & Sitler, R. (1990). Acoustic dimensions of hearing-impaired speakers' intelligibility: Segmental and suprasegmental characteristics. *Journal of Speech and Hearing Research*, *33*, 476–487.

Miller, M. (1984).On the perception of rhythm. *Journal of Phonetics*, *12*(1), 75-83.

Mini, N. (1979). Intonation in Malayalam: Some aspects.Unpublished M.Sc dissertation, University of Mysore.

Mohanan, T. (1989).Syllable structure in Malayalam. *Linguistic Inquiry*, 589-625. Retrieved from: https://www.jstor.org/stable/4178646

Most, T., Harel, T., Shpak, T., & Luntz, M. (2011). Perception of suprasegmental speech features via bimodal stimulation: Cochlear implant on one ear and hearing aid on the other. *Journal of Speech, Language, and Hearing Research*.

Most, T. (1999).Production and Perception of Syllable Stress by Children with Normal Hearing and Children with Hearing Impairment. *Volta Review*, *101*(2), 51-70.

Most, T. (2015).Perception of the prosodic characteristics of spoken language by individuals with hearing loss. *The Oxford Handbook of Deaf Studies in Language*, 79.

Most, T., & Aviner, C. (2009).Auditory, visual, and auditory–visual perception of emotions by individuals with cochlear implants, hearing aids, and normal hearing. *Journal of Deaf Studies and Deaf Education*, *14*(4), 449-464. https://doi.org/10.1093/deafed/enp007

Most, T., & Aviner, C. (2009).Auditory, visual, and auditory–visual perception of emotions by individuals with cochlear implants, hearing aids, and normal hearing. *Journal of Deaf Studies and Deaf Education*, *14*(4), 449-464.

Most, T., & Michaelis, H. (2012).Auditory, visual, and auditory–visual perceptions of emotions by young children with hearing loss versus children with normal hearing. *Journal of Speech, Language, and Hearing Research*, *55*(4), 1148-1162.

Most, T., & Peled, M. (2007). Perception of suprasegmental features of speech by children with cochlear implants and children with hearing aids. *Journal of deaf studies and deaf education*, *12*(3), 350-361. https://doi.org/10.1093/deafed/enm012

Most, T., Bachar, D., & Dromi, E. (2012).Auditory, Visual, and Auditory-Visual Identification of Emotions by Nursery School Children. *Journal of Speech-Language Pathology & Applied Behavior Analysis*, *5(3-4),25-34*.

Murphy, S. T., & Zajonc, R. B. (1993). Affect, cognition, and awareness: affective priming with optimal and suboptimal stimulus exposures. *Journal of personality and social psychology*, *64*(5), 723.

Murray, I. R., & Arnott, J. L. (1993). Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion. *The Journal of the Acoustical Society of America*, *93*(2), 1097-1108. https://doi.org/10.1121/1.405558

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

Nandini, H. M. (1985). Some prosodic aspects in Kannada. Unpublished M.Sc dissertation, University of Mysore.

Narra, M., Teja, D. D., Sneha, M. V. & Dattatreya, T. (2012). Acoustic Correlates of Emphatic Stress in Tulu: A Preliminary Study. American Journal of Linguistics,1 (3), 28-32.

Nataraja, N.P. (1981).Intonation in four Indian languages under five emotional conditions. Journal of All India Institute of Speech and Hearing, 12,22-27.

Nazzi, T., & Ramus, F. (2003).Perception and acquisition of linguistic rhythm by infants. *Speech Communication*, *41*(1), 233-243.

Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human perception and performance*, *24*(3), 756.

Newman, R., Ratner, N. B., Jusczyk, A. M., Jusczyk, P. W., & Dow, K. A. (2006). Infants' early ability to segment the conversational speech signal predicts later language development: a retrospective analysis. *Developmental Psychology.* 42, 643–655. doi: 10.1037/0012-1649.42.4.643

Nittrouer, S., & Chapman, C. (2009). The effects of bilateral electric and bimodal electric—Acoustic stimulation on language development. *Trends in amplification*, *13*(3), 190-205.

O'Halpin, R. (1993). An auditory and acoustic analysis of contrastive stress in profoundly deaf children following training.*Child Language, Teaching and Therapy*, 9, 1.

O'Halpin, R. (2001). Intonation issues in the speech of hearing impaired children:analysis, transcription and remediation. *Clinical Linguistics & Phonetics*, 15, 7,529-550.

Odekar, A. S., (2001). Perception of rhythm in music. An unpublished dissertation submitted to University of Mysore.

Osberger, M. J., & McGarr, N. S. (1982). Speech production characteristics of the hearing impaired. In *Speech and Language* (Vol. 8, pp. 221-283).Elsevier.https://doi.org/10.1016/B978-0-12-608608-9.50013-9

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.

Peters, K. P. (2006). Emotion perception in speech: Discrimination, identification, and the effects of talker and sentence variability (Doctoral Thesis). Retrieved from: https://digitalcommons.wustl.edu/cgi/viewcontent.cgi?referer=https://scholar.google.co.in/&httpsredir=1&article=1112&context=pacs_capstones

Peterson, C. C. (2009). Development of social-cognitive and communication skills in children born deaf. *Scandinavian Journal of Psychology*, *50*(5), 475-483.

Phillips, N. D., Remillard, W., Bass, S., & Pronovost, W. (1968).Teaching of intonation to the deaf by visual pattern matching. *American annals of the deaf*, *113*(2), 239.

Picou, E. M. (2016). How hearing loss and age affect emotional responses to nonspeech sounds. *Journal of Speech, Language, and Hearing Research*, *59*(5), 1233-1246.

Pike K. L. (1946). The intonation of American English.2nd Edition. Ann Arbor: University of Michigan Press.

Pisoni, D. B., & Remez, R. E. (Eds.). (2005). *The handbook of speech perception* (p. 708). Oxford: Blackwell, Victoria 3053, Australia

Polka, L., & Sundara, M. (2003).Word segmentation in monolingual and bilingual infant learners of English and French. In *Proceedings of the 15th international congress of phonetic sciences* (pp. 1021-1024). Retrieved from :https://linguistics.ucla.edu/people/Sundara/pubs/ICPHS_0167.pdf

Pons, F., Lawson, J., Harris, P. L., & De Rosnay, M. (2003). Individual differences in children's emotion understanding: Effects of age and language. *Scandinavian journal of psychology*, *44*(4), 347-353.

Quam, C., & Swingley, D. (2014).Processing of lexical stress cues by young children. *Journal of experimental child psychology*, *123*, 73-89.https://doi.org/10.1016/j.jecp.2014.01.010

Rajupratap, S. (1991). Production of word stress in children - 3 - 4 years. In M. Jayaram & S.R.Savithri (Eds).Research at AIISH.Dissertation.Abstracts-3, 11.

Ramkalawan, T. W., & Davis, A. C. (1992).The effects of hearing loss and age of intervention on some language metrics in young hearing-impaired children. *British Journal of Audiology*, *26*(2), 97-107.

Ramus, F., Hauser, M. D., Miller, C., Morris, D., & Mehler, J. (2000).Language discrimination by human newborns and by cotton-top tamarin monkeys. *Science*, *288*(5464), 349-351.

Raphael, L. J., Borden, G. J., & Harris, K. S. (2007). *Speech science primer: Physiology, acoustics, and perception of speech*. Lippincott Williams & Wilkins, Walnut street, Philadelphia.

Ravisankar, G. (1987). Intonation of Tamil.Ph.D thesis, Annamalai University.

Rigault, A. (1962). Role of vowel frequency, intensity and duration in the perception of the accent in French. In *Proceedings of the Fourth International Congress of Phonetic Sciences, Helsinki* (pp. 735-48).

Rosenhouse, J. (1986). Intonation problems of hearing-impaired Hebrew-speaking children. *Language and speech*, *29*(1), 69-92.

Rubin-Spitz, J., McGarr, N. S., & Youdelman, K. (1986). Perception of stress contrasts by the hearing impaired. *The Journal of the Acoustical Society of America*, *79*(S1), S10-S10.https://doi.org/10.1121/1.2023065

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

Savithri, S. R. (1999b). Perceptual cues of word stress in Kannada. The Journal of the Acoustical Society of India, 25, 1-4. Retrieved from: http://drsavithri.com/articles/Perception%20of%20word%20stress%20in%20Kannada.pdf

Savithri, S. R., Agarwal, R., Rani, R., & Manasagangothri, M. (2008). SPEECH RHYTHM IN HEARING-IMPAIRED. *Proceedings of FRSM*, *2008*.

Savithri, S. R., Jayaram, M., Kedarnath, D. & Goswami, S. (2006). Speech rhythm in Indo Aryan and Dravidian languages.*Proceedings of the International Symposium on Frontiers of Research on speech and music*, 31-35.

Savithri, S. R., Maharani, S., Sanjay, G. & Deepa, D. (2007). Speech Rhythm in Indian Languages.*AIISH Research Fund Project*.

Savithri, S. R., Johnsirani, R. & Ruchi, A. (2008).Speech Rhythm in Hearing-Impaired Children.*AIISH Research Fund Project*.

Savithri, S. R., Sreedevi, N., Jayakumar, T., & Kavya, V. (2010).DEVELOPMENT OF SPEECH RHYTHM IN KANNADA SPEAKING CHILDREN. *Journal of the All India Institute of Speech & Hearing*, *29*(2).

Savithri, S.R., Sreedevi, N., Deepa Anand., & Aparna.V.S. (2011). Effect of gender on speech rhythm in 3-4 year old Kannada speaking children.*Proceedings of International Symposium on Frontiers of Research in Speech and Music & Computer Music Modeling and Retrieval, FRSM/CMMR-2011,* 31-35.

Scherer, K. R. (2003). Vocal communication of emotion: A review of research paradigms. *Speech communication*, *40*(1-2), 227-256.https://doi.org/10.1016/S0167-6393(02)00084-5

Schorr, E. A., (2005). *Social and Emotional Functioning of Children With Cochlear Implants* (dissertation). CollegePark: University of Maryland.

Scott, D. R., Isard, S. D., & de Boysson-Bardies, B. (1985).Perceptual isochrony in English and in French. *Journal of Phonetics*, *13*(2), 155-162.

Seidl, A., & Cristià, A. (2008). Developmental changes in the weighting of prosodic cues. *Developmental Science*, *11*(4), 596-606.

Seidl, A., & Johnson, E. K. (2008). Boundary alignment enables 11-month-olds to segment vowel initial words from speech. *Journal of child language*, *35*(1), 1-24.

Sethi, J. (1971). Intonation of statements and questions in Punjabi. Hyderabad: CIEFL

Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science*, *270*(5234), 303-304.

Shinall, A. R. (2005). Emotion perception in pre-kindergarten school children at Central Institute for the Deaf (Master's Thesis).*Retrieved from* //digitalcommons.wustl.edu/cgi/viewcontent.cgi?article=1438&context=pacs_capstones

Snow, D. (1998). Children's imitations of intonation contours: Are rising tones more difficult than falling tones?. *Journal of Speech, Language, and Hearing Research*, *41*(3), 576-587.

Sorenson, J. (1974). Fundamental frequency contours in the speech of deaf children. Unpublished manuscript, Massachusetts Institute of Technology

Spring, D. R., & Dale, P. S. (1977).Discrimination of linguistic stress in early infancy. *Journal of Speech and Hearing Research*, *20*(2), 224-232. https://doi.org/10.1044/jshr.2002.224

Srinivas, C. (1992). Word stress in Telugu and English. Unpublished Dissertation submitted to C.I.E.F.L.

Stark, R. E., & Levitt, H. (1974).Prosodic feature reception and production in deaf children. *The Journal of the Acoustical Society of America*, *55*(S1), S63-S63.https://doi.org/10.1121/1.1919842

Swingley D. (2009) Contributions of infant word learning to language development.PhilosTransRSocLondBBiolSci 364(1536):3617–3632.

Thiessen, E. D., Hill, E. A., & Saffran, J. R. (2005). Infant-directed speech facilitates word segmentation. *Infancy*, *7*(1), 53-71.

Thiessen, E. D., & Saffran, J. R. (2003).When cues collide: Use of statistical and stress cues to word boundaries by 7- and 9-month-old infants. Developmental Psychology, 39, 706–716.

Thiessen, E. D., & Saffran, J. R. (2007).Learning to learn: Infants' acquisition of stress-based strategies for word segmentation. *Language learning and development*, *3*(1), 73-100.https://doi.org/10.1080/15475440709337001

Toe, D. M., & Paatsch, L. E. (2013).The conversational skills of school-aged children with cochlear implants. *Cochlear implants international*, *14*(2), 67-79.

Tooher R. (2001) Relationships among speech perception, production, language, hearing loss, and age in children with impaired hearing. J Speech Lang Hear Res 44(2):264–285.

Trager, G. L., & Smith, H. L. (2009). *An outline of English structure.Activities for practicing stress and intonation.English Teaching Forum* (Vol. 20, pp. 15-18).

Tye-Murray, N. (2003). Conversational fluency of children who use cochlear implants. *Ear and Hearing*, *24*(1), 82S-89S.

Van Wieringen, A., & Wouters, J. (2015). What can we expect of normally-developing children implanted at a young age with respect to their auditory, linguistic and cognitive skills?. *Hearing research*, *322*, 171-179.

Vavatzanidis, N. K., Mürbe, D., Friederici, A. D., & Hahne, A. (2016). The perception of stress pattern in young cochlear implanted children: an EEG study. *Frontiers in neuroscience*, *10*, 68.https://doi.org/10.3389/fnins.2016.00068

Voelker, C. H. (1935). A preliminary strobophotoscopic study of the speech of the deaf. *American Annals of the Deaf*, 243-259.

Werker, J. F., & Yeung, H. H. (2005). Infant speech perception bootstraps word learning. *Trends in cognitive sciences*, *9*(11), 519-527.

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.https://doi.org/10.1177/002383096600900205

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

Yildirim, S., Bulut, M., Lee, C. M., Kazemzadeh, A., Deng, Z., Lee, S.,.& Busso, C. (2004). An acoustic study of emotions expressed in speech. In *Eighth International Conference on Spoken Language Processing* (pp. 2193-2196).Retrieved from: https://www.isca-speech.org/archive/archive_papers/interspeech_2004/i04_2193.pdf

Yoshinaga-Itano, C. (2003). From screening to early identification and intervention: Discovering predictors to successful outcomes for children with significant hearing loss. *Journal of deaf studies and deaf education*, *8*(1), 11-30.

Yoshinaga-Itano, C., Sedey, A. L., Coulter, D. K., & Mehl, A. L. (1998).Language of early-and later-identified children with hearing loss. *Pediatrics-English Edition*, *102*(5), 1161-1171.

Ziv, M., Most, T., & Cohen, S. (2013). Understanding of emotions and false beliefs among hearing children versus deaf children. *Journal of Deaf Studies and Deaf Education*, *18*(2), 161-174.