COMPARISION OF SYLLABIC AND DUAL COMPRESSION ON NON-LINEAR HEARING AID PROCESSED MALAYALAM CHIMERIC SENTENCES

UDHAYAKUMAR R
Register Number: 16AUD032

This Dissertation is submitted as part fulfilment

For the Degree of Master of Science in Audiology

University of Mysuru, Mysuru

ALL INDIA INSTITUTE OF SPEECH AND HEARING

MANASAGANGOTHRI, MYSURU-570006

MAY 2018

CERTIFICATE

This is to certify that this dissertation entitled **"Comparison of syllabic and dual compression on non-linear hearing aid processed Malayalam chimeric sentences"** is the bonafide work submitted in part fulfilment for the degree of Master of Science (Audiology) of the student Registration Number: 16AUD032. This has been carried out under the guidance of the faculty of the institute and has not been submitted earlier to any other University for the award of any other Diploma or Degree.

Mysore
May, 2018

Prof. S. R. Savithri
Director
All India Institute of Speech and Hearing
Manasagangothri, Mysore-570006

CERTIFICATE

This is to certify that this dissertation entitled **"Comparison of syllabic and dual compression on non-linear hearing aid processed Malayalam chimeric sentences"** has been prepared under my supervision and guidance. It is also being certified that this dissertation has not been submitted earlier to any other University for the award of any other Diploma or Degree.

Mysuru
May 2018

Dr. Devi. N
Guide
Reader in Audiology
All India Institute of Speech and Hearing
Manasagangothri,
Mysuru-570006

DECLARATION

This is to certify that this dissertation entitled **"Comparison of syllabic and dual compression on non-linear hearing aid processed Malayalam chimeric sentences"** is the result of my own study under the guidance of Dr Devi N, Reader in Audiology, Department of Audiology, All India Institute of Speech and Hearing, Mysore, and has not been submitted earlier to any other University for the award of any other Diploma or Degree.

Mysore **Registration No: 16AUD032**

May, 2018

Dedication

என் தாய்க்கும் தந்தைக்கும், என்
பிறப்பில் இருந்து இதுவரை
எனக்கு இன்னொரு தாய் ஆகவும்
தந்தை ஆகவும் நல்ல
நண்பனாகவும் இருக்கின்ற என்
அண்ணன்களுக்கும், எனக்கு
பக்கபலமாக இருக்கும்
நண்பர்களுக்கும் என் நலன்
கருதும் நல்ல உள்ளத்திற்கும்
மற்றும் என் சிறந்த
வழிகாட்டியான டாக்டர். தேவி
அவர்களும் இந்த படைப்பினை
சமர்ப்பிக்கிறேன்

Acknowledgements

Completion of my dissertation won't be possible without several people. At this moment I would like to extend my sincere gratitude. First of all I am grateful to sincerely thank my dissertation guide Dr. Devi N, Reader in Audiology, All India Institute of Speech and Hearing, Mysuru for her constant support, motivation, scholarly input and her valuable guidance throughout the completion of my dissertation and add on to it I never saw such a coolest guide ever thank you so much ma'am for making me complete my dissertation without any stress throughout my completion of my dissertation work. Thank you so much ma'am.

I thank Dr. S R Savithri, Director AIISH, Mysuru for permitting me to do my study. I thank HOD, Department of Audiology for providing academic support and facilities provided to carry out the research work at the institute.

I would like to acknowledge several faculty members of this institute for their kindly help throughout several phases of my study by extending their knowledge. Firstly I would like to extend my gratitude to Mr. Nike Gnanateja, Mr Anoop, and Dr. Sharathkumar *for helping me during my stimulus preparation and recording. Dr. Vasanthalakshmi, Reader in Bio statistics for her support in statistical analysis and explaining the results. I sincerely thank you ma'am.*

I would like to sincerely extend my gratitude to Sneha akka & Kriti akka for spending their valuable time during the preparation of stimulus, all the participants of my study, thank you all for your valuable time spent during my dissertation work.

I would like to thank all my friends: BESTIES FOREVER, CORE ROCKERS, SUSTAINERS, FIREBOLT and the awesome ones the KWATLE BOYS for all their support throughout my dissertation work.

My joyous gratitude to my care taking good souls Charming Revathi Aunty, Awesome Prabu Uncle, Sottai Vaishu matha, Kulachi Lavy, Machan's for life (Koraati Aswin, Kannadi Arun, Sidhar Sivaras, Poli Doctor Saravanan), Kuttachi Priya, Rajubhai Mohan anna, Crazy Saji akka, Advisor Ram anna, Anime Meenu akka, Thalaivar Sathiya anna, Chocolate boy Jagacharan, Chockie Pie Shubha, Lepotig Anu, Caring Ankita, Silent Sanjeev anna, Free TV Prithivi,

Pathiyam Vasu, Thani Shiyaam, Dr Napoleon anna, Yasin Bhai, Manjunath (or) Monkey man Chandu, Gym boy Vishwa, Thiviravathi chandan, Kwacha Rakshit, Hari….sha, Dumdare suju, Bule Anup, NA Ethe, Pataka Pathak, 2.5 Sam, Director Mangal, Fruitboy Nishanth, Kuppa Teena, Dancer Sai Shruthi.

The last but not the least the girl who bares me all the time, shouts at me, care at me, beside me, be there with me in all situation, scold me the one and only my dear Doly(Ranjitha).

I owe a lot to my parents, who encouraged and helped me at every stage of my personal and academic life and longed to see this achievement come true. I am much indebted to my family who supported me in every possible way to see the completion of this work.

Abstract

The human auditory system perceives any speech sound through the inherent temporal cues i.e., the temporal fine structure (TFS) and envelope (ENV) cues which has its own predominance for perception across languages. Research in English and Mandarin Chinese language showed the difference between these cues with tonal language employ more of TFS cues and non-tonal language employ ENV cues for perception. Earlier studies on Indian language (Kannada and Malayalam) revealed ENV cues predominance for perception. Based on this the aim of this study was to compare the syllabic versus dual compression system on a nonlinear hearing aid processed stimuli using Malayalam chimeric sentences. Thirty-five normal hearing individual were assessed for the perception of the chimeric sentence across eight frequency bands from 1 to 64 bands. The result of the present study revealed there is a significant difference across frequency bands between two dual and syllabic compressions. The ENV cues perceived better from the 4 frequency bands with TFS cue predominant at 1 frequency band and also across compressions. The syllabic compression showed a marginal difference between the compressions.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

The human auditory system has the ability to hear most complex sound signals due to a wide range of frequency of hearing with lower to upper limit of hearing at 20 Hz to 20 kHz. The human auditory system is a complex structure constituted by outer, middle and inner ear that helps in collecting, transmitting and transducing the sound signal to the higher structures for further processing. The sound reaches the outer ear in the form of acoustic energy which will be converted into mechanical energy at the level of middle ear due to the transfer action which further transfer this into the cochlea where it behaves as a frequency analyzer as a support from place theory of hearing as each place on the cochlea responding more favorably to a particular frequency known as the characteristic frequency (CF) (Walsh, 1960). The sensory hair cells in the cochlea transduct the mechanical energy into neural spikes at the level of Auditory Nerve (AN) fibres innervating the hair cells. The neural spikes at the AN fibers is coded and the neural code is sent to the higher structures through the higher order neurons for further analysis and processing of signals where the auditory system performs several tasks such as sound localization, speech understanding, pitch and melody perception, which are usually required to function properly even in the existence of competing for speech. Understanding the mechanism of operation of the auditory system requires good attention to the structure of the different sections in the auditory system as well as the interactions between the different parts (Billone & Ragnar, 1973).

Intensive research has been going on in the area of speech perception to identify the factors and mechanism by which humans understand speech in different listening conditions. When a sound is received by the cochlea, the frequency content of the signal is mapped into a pattern of excitation along the Basilar Membrane called the auditory filters (Chen, Clark, & Jones, 2003). In normal ears, the auditory filters were sharply tuned however in the case of the impaired ear the filters were broadly tuned (Florentine, Buus, Scharf, & Zwicker, 1980; Moore, Glasberg, & Stone, 1999). An acoustic signal entering the ear gets bandpass filtered into individual frequency components in congruence with the basilar membrane excitation on reaching the sensory end organ. Thus, these structures are involved in simple frequency segregation to complex analysis of segregating acoustic signal from background noise. Excitation patterns code the spectrum information of the acoustic stimulus, which is referred to as 'spectral' or 'place' information (Nuttall, Christian Brown, Masta, & Lawrence, 1981). It is assumed that spectral information plays an important role in speech recognition as many phonetic features are characterized by their frequency spectrum. Because basilar membrane acts as a bank of bandpass filters, it yields in the frequency selectivity. Analyzed signals consist of information in two major forms i.e. slow varying amplitude fluctuations over time known as the envelope (E) and rapid oscillations with a rate close to the Centre frequency of the band known as temporal fine structure (TFS) (Moore & Sęk, 2009; Rosen, 1992). Usually, the TFS is denoted as the "carrier" whereas the (E) acts as the "amplitude modulator" to the carrier. In the auditory system, these E cues are represented as the short term firing of auditory neurons, while TFS are denoted by the synchronous firing to the phase of the carrier i.e. phase locking. (Joris & Yin, 1992; Rose, Brugge, Anderson, & Hind, 1967)

During the processing of the acoustic signal, the envelope cues and fine structure cues have its own relative contribution in identifying the signal. However, special algorithms are required in order to understand individual contribution of each cue in the auditory processing of the signal. In order to investigate the relative role of both envelop and fine structure, earlier studies have used method of peak clipping where there is an increase in impairment of speech intelligibility (Licklider & Pollack, 1948) and Hilbert transform method (Bracewell & Bracewell, 1986) to separate the TFS and E cues in a signal to study perceptual significance of the E and TFS cues by constructing hybrid sound known as the 'auditory chimeras'(Soli et al., 1998). Chimeric sentences are developed by interchanging the TFS and E cues of two different sentences and hence help us in understanding the importance of perceiving each sound. Perceptual studies on normal auditory system demonstrate that envelop cues are sufficient for speech perception in quiet, whereas fine structure is important in pitch perception, lexical tone perception and also speech perception in noisy conditions (Loizoumichael, Tu, & Dorman, 2006; Smith, Delgutte, & Oxenham, 2002). The TFS information is coded through the phase locking property of the AN fibres and it is known that phase locking is weak at high frequencies with almost a complete loss of synchrony for frequencies above 4-5 kHz in mammalian auditory systems (Palmer & Russell, 1986). Hence, it is commonly assumed that 'TFS information is not used for frequencies above that limit'. It has been demonstrated in many experiments that envelope (ENV/E) cue is necessary for speech understanding and it provides robust speech recognition in silence even when provided in as few as four frequency bands (Flanagan, 1980; Shannon, Zeng, Kamath, Wygonski, & Ekelid, 1995;

Smith et al., 2002). Hence auditory chimaeras provide a way to study the relative importance of envelope and fine structure in speech perception and pitch perception.

Smith et al., (2002) found that in speech–speech chimaeras with 4-16 frequency bands, the words represented in the envelope were identified correctly much more frequently than words represented in the fine structure. However in the same study melody chimeras were prepared by extracting the envelope from one melody and fine structure from other melody and combined together across 9 frequency band the participants were given a pitch perception task from the prepared complex harmonic stimuli and all the participants were reported of perceiving the temporal fine structure cue till 32 frequency bands and the perception envelope cues were predominant only at 48 and 64 frequency bands being the cutoff higher for melody chimeras when compared to speech-speech chimeras, which concludes the importance of specific cues and the number of frequency bands needed to identify speech and melody.

Auditory abilities and speech perception are known to mature over the first 10–12 years (Hnath-Chisolm, Laipply, & Boothroyd, 1998; Siegenthaler, 1969). Hence, study done by Eisenberg, Shannon, Martinez, Wygonski, and Boothroyd., (2000) to rule out the developmental trend in children age ranging from 5 to 12 years on maturation of ENV and TFS cues perception using noise-vocoded speech stimuli revealed that the ability of using E cues start to mature at the age of 7 years before and attains adult like maturation around 10 years of age.

The temporal auditory mechanisms involved in the perception of phonetic features are crucial for language acquisition even in infants (Nazzi, Iakimova, Bertoncini, Frédonie,

& Alcantara, 2006; Nazzi, Jusczyk, & Johnson, 2000). Research on the individual with hearing impairment with cochlear origin on adult participants revealed that the cochlear hearing loss deteriorates the ability to encode TFS cue by preserving the ability to encode ENV cues (Gilbert, Lorenzi, Gilbert, & Lorenzi, 2006). Impaired perception of envelope cues has also been reported repeatedly in the case of central damage to the auditory system and specific language acquisition disorders (Giraud et al., 2000; Lorenzi, Dumont, & Fu¨llgrabe, 2000)

Research work on individual with sensorineural hearing loss on perception of TFS cues using speech recognition task revealed a difficulty in perception of TFS cue which in turn supported by psychoacoustic measures that individual with hearing impairment showed an impaired phase-locking ability, might be a possible reason for poor encoding of TFS cues (Buss, Hall, & Grose, 2004). Researchers used the Hilbert transform for separating the E and TFS cues and presented these stimuli to the hearing impaired individual and checked for the perception of speech across three conditions such as the unprocessed, TFS and E condition with training the perception of unprocessed and E cue only perception was found to be better, however, the hearing impaired individual exhibited poorer score under TFS only condition because of their impaired ability of processing the TFS cue in hearing impaired individual. Hence, these individual uses more amount of the E cues compared to TFS cues for the perception of speech.(Moore & Lorenzi, 2006). Hence, the present study was to investigate the effect of hearing aid processed stimuli on the perception of the envelope and fine structure cue across frequency bands in the Malayalam language.

Need for the study

Heinz and Swaminathan, (2009) studied the effect of the envelope and fine structure cues in normal hearing individuals especially in foreign tonal and non-tonal languages (Mandarin Chinese & English) and found that in tonal languages TFS plays important role in the perception (Xu, Thompson, & Pfingst, 2005). Smith et al., (2002) found that the non-tonal language was governed by the envelope (E) cues however tonal language was governed by the temporal fine structure (TFS) cues in the normal hearing individual.

Studies were done on south Indian languages in Malayalam using Malayalam chimeric sentence and in Kannada using Kannada chimeric words and sentences. In Malayalam, the speech stimuli with the lower bands were identified with TFS cues whereas speech stimulus with the higher bands used E cues (Indu & Devi, 2016). However, when Kannada chimeric words and sentences were used, the results revealed that envelope cues were employed majorly than fine structure cues (Naveen & Devi, 2017). As divergent results have been reported in different languages there is a need to investigate the relative importance of cues in each of these languages separately. In case of individuals with hearing loss of cochlear origin studies had revealed that the individual with hearing loss had difficulty in encoding TFS cues because of the impaired phase-locking abilities. Buss, Iii, and Grose., (2004) also supported that individuals with the sensorineural hearing loss had reduced ability in encoding TFS cues due to impaired phase locking. Hopkins, Moore, and Stone., (2008) had concluded that an individual with moderate cochlear hearing loss has reduced the ability to use the TFS cues to understand speech both in the quiet and noisy situation. The need of utilizing the syllabic compression in the present is that it is a fast acting compression system which is more helpful for an individual with hearing loss (Dreschler, 1989).

Automatic volume control (AVC), can theoretically benefit people with hearing loss who do not have a severely reduced dynamic range but who have an optimum level for speech discrimination above which their performance starts to decline (King & Martin, 1984). The researcher also asserts that AVC allows the listener to accept a wider range of input levels, without continually adjusting the gain that would be possible with a linear aid.

Dual compression is a combination of both syllabic and AVC. It exploits the advantage of both syllabic compression and AVC (Moore & Glassberg, 1991). Syllabic compression and dual compression had the same effect on speech identification scores (SIS) for both quiet and in noisy conditions but the hearing aid user preferred to use dual compression (Geetha & Manjula, 2005)

Hence, processing of chimeric sentences in hearing aid with syllabic and dual compression and assessing it on normal hearing individual provides us with a better understanding of how temporal cues be perceived through hearing aid in individuals with hearing impaired. In this present study, the hearing aid processed speech sentences with syllabic and dual compression was tested in normal hearing individuals.

Aim

The aim of the study was to compare the effect of syllabic compression and dual compression on nonlinear hearing aid processed Malayalam chimeric sentences in normal hearing individuals.

Objective:

- To study the influence of envelope cue on hearing aid processed Malayalam chimeric sentences using syllabic compression and dual compression
- To study the influence of fine structure cue on hearing aid processed Malayalam chimeric sentences using syllabic compression and dual compression
- To compare between the two compressions in the hearing aid
- To find out the number of frequency bands required to clearly differentiate between the envelope and fine structure cues

Hypothesis:

The following null hypotheses were framed for each main objective of the study. They were:

- There is no significant effect of envelope cue on hearing aid processed Malayalam chimeric sentences using syllabic compression and dual compression
- There is no significant effect of fine structure on hearing aid processed Malayalam chimeric sentences using syllabic compression and dual compression
- There is no significant effect between the two compressions in the hearing aid.
- There is no significant effect of identification of chimeric sentences across the frequency bands.

Chapter 2

Review of literature

If a sound reaches the cochlea at the level of the basilar membrane the frequency content of it is mapped along the length of the cochlea based on the excitation pattern. This excitation pattern code the spectral or the place information of the acoustic information which is concerned (Dallos, Popper, Fay, & Dallos, 1996). This spectral information plays a major role in recognition of language with the help of the phonetic features characterized by the frequency spectrum of the stimulus. The basilar membrane is regarded as the bank of frequency filters corresponds to the place along the cochlear length which in turn provides the frequency selectivity of any stimulus (Palmer & Russell, 1986). The output of these bandpass filters provides the important temporal information which was the fast oscillations or the temporal fine structure cues (TFS) superimposed by a gradually varying temporal envelope cue (ENV). This temporal information was further provided to the sensory nerve (Smith et al., 2002; Xu & Pfingst, 2003; Zeng et al., 2004).

2.1 Envelope versus temporal fine structure cues for speech perception and predominance across languages

The change in the firing rate at the level of Auditory nerve (AN) fibres are associated to the envelope of the signal and the time between the spikes reflects the temporal fine structure or TFS (Young & Sachs, 1979). The relative envelope magnitude across channels carries information that can be used in the auditory system to identify the signal spectral shape and its slow short-term spectral changes. The TFS conveys cues regarding the central frequency of the sound and about its short-term spectrum.

The phase locking property of the auditory nerve (AN) fibres help in coding the TFS information, however, this property decreases as the frequency increases i.e., the phase locking is best at the low frequency and gets weaker or absent above 4-5 kHz for mammalians (Palmer & Russell, 1986). Hence, it is to be believed that the TFS information is not used above this frequency boundary instead for robust speech identification in the quiet environment the envelope or ENV information usually assumed to be significant in understanding speech with as few as four frequency bands (Flanagan, 1980; Shannon et al., 1995; Smith et al., 2002). Whereas in case of the noisy situation this four bands is not sufficient for understanding speech using the ENV cues the individual requires more number of frequency bands of ENV cues to perceive speech (Qin & Oxenham, 2005; Stone & Moore, 2003).

There are several studies which has been carried out in identifying the predominance of ENV vs TFS cues on the perception of speech. A study by Fu, Zeng, Shannon, Soli, and Shannon., (2006); Pfingst, Arbor, and Pfingst., (2003); Smith et al., (2002) in which the comparative role of speech envelope (E) and temporal fine structure (TFS) cues on speech identification was assessed in which one cue is varied while the other was kept intact this was accomplished by the vocoders. This vocoded speech is generated by extracting the two cues from the speech signal and the envelope cue is set to modulate the tone or noise by separating it across frequency bands and the importance of this envelope cues was studied. However, there are studies examined the potential contribution of TFS cue also.

Xu and Zheng, (2007) assessed the comparative assistance of spectral and temporal cues to phoneme recognition, in which the spectral information was altered by changing the channels in the vocoder processing, however, the temporal cues were altered by changing the cutoff frequency of the envelope (E) using the low pass filter. The recognition of vowel and consonant using this method was studied and found an existence of a tradeoff between the spectral and temporal cues in identification which is because of enhancement of anyone cue reduces the other cue perception.

Nie, Barco, and Zeng., (2006), reported on normal hearing individual and individuals with hearing impairment on identification of predominant cue for speech perception for that the spectral and temporal cues. The cues were altered by varying the number of channels and pulse rate, which revealed that the TFS cues were efficiently used by the normal hearing individual when compared with the hearing impaired. Supporting this findings, others Hopkins and Moore., (2007, 2009); Lorenzi, Gilbert, Carn, Garnier, and Moore., (2006); Brian C. J. Moore and Sęk., (2009); Brian C.J. Moore, Glasberg, and Hopkins., (2006) have also reported that it is due to the impaired ability in individuals with hearing impairment on their phase locking ability, which in turn reduced their ability to decode the fine structure information. Lastly, due to the broad auditory filters will also be a substantial evidence to portrait the poor performance on understanding the fast changing TFS information in hearing impaired individual (Moore & Glasberg, 1987; Stone, Füllgrabe, & Moore, 2008).

In accordance to the earlier researchers a study by Hopkins and Moore., (2007, 2009) on normal and hearing-impaired individual where the benefit of TFS cue was assessed using a sentence recognition task in which the threshold was obtained i.e., sentence recognition threshold (SRT), is the lowest hearing level in which an individual can identify 50% of the speech material. The researchers have quantified the importance of the TFS cue by altering the number of frequency bands containing the TFS information on a tone or noise vocoded stimuli. To examine the known fact that the hearing impaired individual had poor ability to use the TFS cue at high frequency in this study the researchers removed the TFS cues across the low, mid and high-frequency range and found that the hearing impaired individual had poor ability to use TFS at the mid and high-frequency range.

Hopkins and Moore, (2009) measured the SRTs in normal hearing subjects while varying the cutoff channel which was the frequency band below which the stimulus was left intact, while TFS information was deleted from all bands above it. It was established that the SRT declined significantly as the value of the cutoff channel increased, which suggests that TFS has a significant part in understanding speech in fluctuating background noise.

Gnansia, Péan, Meyer, and Lorenzi., (2009) studied the effects of spectral smearing and degradation of TFS cues on masking release, which is the capability to listen in the dips of the background noise. Stimuli were processed using a spectral smearing algorithm or a tone vocoder technique. The spectral smearing algorithm computes the short-term spectrum using Fast Fourier transforms (FFT), and then the spectrum is smeared by a divisor of two or four using a smearing matrix for 2-ERBN or 4-ERBN auditory filters. The fundamental frequency information was more degraded by the vocoder than the spectral smearing algorithm. Masking release was reduced more with the tone vocoder than spectral smearing. The conclusion of the study showed that both frequency selectivity and TFS cues are significant for the ability to listen in the dips. Lorenzi et al., (2006) assessed the comparative use of ENV and TFS cues in reconstructing missing information in interrupted speech. In their subject field, four types of sentences sort out into 32 frequency bands and information in 21 bands were removed or sort out so that the final stimuli have different amounts of ENV and TFS cues. Four types of sentences were used; reference, partially empty, vocoded and partially vocoded. The resulting sentences were still understandable but the intelligibility significantly deteriorated after adding a silence gap. TFS cues have a significant part in reconstructing the broken sentences. The TFS is not sufficient alone, but when the individuals were trained with ENV cues the perception of interrupted speech increased.

A significant concern regarding the results for TFS contribution to speech understanding is that these effects may be influenced by possible ENV cues in signals. These ENV cues may be due to inefficient signal processing techniques applied to separate TFS from ENV, as TFS and ENV are not totally independent (Ghitza, 2001). Some other significant constituent was that the possible recovery of ENV cues by the human auditory filters from a correctly processed signal having only TFS cues. For example, narrow-band filtering can recover the signal ENV from the fine- structure information (Voelcker, 1966). This is especially important in humans because of the sharp cochlear tuning (narrow filters), which facilitate the retrieval of the slow amplitude variations (ENV) from the TFS signal (Ghitza, 2001; Heinz & Swaminathan, 2009; Zeng et al., 2004). Gilbert et al., (2006) argued that recovery of ENV cues from TFS-only signals has minimal contribution to speech recognition when the vocoder analysis filters, which are utilized to generate the TFS-only stimulus, have a bandwidth less than 4 ERBN. According to the researchers, using 16 frequency channels should be sufficient to prevent the use of recovered ENV cues. Heinz and Swaminathan, (2009), nevertheless, presented physiological evidence for the presence of recovered ENV in chinchilla Auditory Nerve response to the chimeric speech. It was computed that 'Neural cross-correlation coefficients' measures the similarity between ENV and TFS to quantify the similarity between ENV or TFS components in the spike train responses.

Sheft, Ardoint, Lorenzi, and Sheft (2008) presented different ways to reduce the fidelity of ENV reconstruction from TFS signals. The TFS signal can be filtered by an all-pass filter with a random phase response. The results were on the assumption that ENV and the instantaneous phase are connected, so that processing the TFS signal to produce a mismatch with the original ENV signal would reduce the fidelity of ENV recovery (Schimmel & Atlas, 2005). The other method to reduce the chances of meaningful ENV recovery from TFS cues was to increase the number of analysis filters. When the bandwidth of the analysis filter are narrower than 4 times the normal auditory filter, some studies argued that the role of recovered ENV cues in speech perception is negligible (Gilbert et al., 2006)

The last method proposed by Sheft et al., (2008) was to limit the bandwidth of the extracted TFS signal of the analysis filter bandwidth in order to degrade ENV reconstruction. The results of revealed that TFS stimuli, processed to reduce chances of intelligibility from recovered ENV cues, were still highly intelligible (50%– 80% correct consonant identification).

Indu and Devi, (2016) investigated the perception of ENV versus TFS cues using speech identification task in Malayalam language using Malayalam chimeric sentences across 8 frequency bands and found that the predominance of TFS cues are more at lesser frequency band, as the number of frequency bands increased the predominance in cue were shifted towards the envelope of the sentence for perception and hence concurrent to this finding Naveen and Devi, (2017) assessed the same in Kannada language using Kannada chimeric words and sentences and the results revealed that for words the predominance of TFS cues persist till 16 bands however in case of sentence the trend is followed similar to that of the earlier study on Malayalam language being at increase in bands the perception of ENV cues was higher.

2.2 Processed stimuli on recognition of speech identification

Lorenzi et al., (2006) evaluated the contribution of TFS in speech identification, consonant identification in nonsense processed vowel-consonant-vowel (VCV) stimuli were used. The individuals with hearing impairment and normal subjects were tested with TFS-only stimuli generated from nonsense VCV words by separating out the original signal into 16 contiguous frequency bands, computing the E and FS in every band using the Hilbert transform, and combining the TFS signals from the different frequency bands to construct the last stimulus. Results revealed that normal hearing subjects show significant intelligibility for TFS cues, where up to 90% recognition is reported after some training. Hopkins et al., (2008) explained the need for training to achieve high recognition scores by the possibility that the auditory system is not applied for processing TFS cues in isolation from ENV cues or that TFS cues in processed stimuli are distorted compared to intact speech and hence training are needed. In a similar experiment Lorenzi et al., (2010), it has been shown that children with normal hearing aged 5 to 7 are able to make use of TFS cues. Results revealed that normal hearing children can use both ENV and TFS cues at the same level as adults, which means that tests for the sensitivity to TFS cues can be performed at this very youthful age for the early spotting of any potential problems in the TFS process.

Hopkins and Moore., (2009) measured the SRTs for speech processed to contain varying amounts of TFS cues. The speech signals were filtered using 30 1-ERBN filters and treated to keep ENV only information or left unprocessed to preserve both ENV and TFS cues. It was noticed that when there are more channels containing TFS cues, SRT was decreased, showing benefits from the introduced cues. Redundancy in TFS information as adding TFS in some channels does not always improve the threshold was also observed. Another experiment was carried out where filtering of the speech signal through 56-ERBN channels and brought forth a tone vocoded signal in four of the available five channels. The fifth channel was either absent or was unprocessed. Normal hearing subjects benefitted from the added TFS cues over a spacious range of frequency, while the benefit was less in hearing-impaired subjects.

2.3 Compressions (Dual Vs Syllabic) in hearing aids - its effect and benefits on speech perception

The compression system of a hearing aid basically works on the duration of the attack time and release time-based on this there is two compression system that is the fast acting (syllabic) and slow acting (dual) compression system. The Fast acting compression has a relatively short attack and release time to amplify and sound signal such as the words or syllables (Walker, Byrne, & Dillon, 1984). This time constant of a fast acting system are typically set at duration to shorten then the typical syllable i.e., approximately 200 to 300 ms (Hickson, 1994). The frequency-gain shape characteristics also change rapidly based on the short-term speech spectrum of the incoming signal (Moore, Peters, & Stone, 1999). This type of compression system felicitates an increase to a low-level signal which induces an upward spread of masking of a low-level signal (Byrne, Dillon, Ching, Katsch, & Keidser, 2001). Fast acting compression provides an improvement in speech perception in noise because it helps in perceiving speech from spectral and temporal dips of noise though it does not replace normal performance (Moore et al., 1999). There are contradictory opinions also to syllabic compression. Benson, Clark, and Johnson., (1992) in their research work state that the syllabic compression reduces the peak to valley ratio of any ongoing speech or noise to the hearing aid. This reduction in the peak and valleys due to the syllabic compression system to a speech signal cause distortion (Kuk, 1999).

Slow acting compression are those which has a longer attack and release time for compression. The time constants are usually longer than a typical syllable or a word. Usually, release time can vary from 150 ms to several seconds which means that system adjusts the gain by considering the long-term spectrum of the incoming signal (Hickson, 1994).

Several research works had been carried out to evaluate the effect of the dual or slow-acting compression system on speech identification showed variation in results. Neuman, Bakke, Mackersie, Hellman, and Levitt., (1995) evaluated the effect of release time in compression hearing aids revealed that the participants preferred longer release time of around 200 and 1000ms to perceive a speech in presence of high-level noises like cafeteria noise. But this longer release time would result in loss of information in a situation where the level of the incoming signal changes rapidly (Byrne et al., 2001)

Geetha and Manjula., (2005) studied the effect of syllabic and dual compression on speech perception in quiet and in noise using a non-linear hearing aid programmed with syllabic and dual compression. Speech identification scores were obtained in sound field at two different presentation levels (45 dB HL and 70 dB HL), in quiet and in the presence of noise (SNR + 10 dB). Results showed that even though there was no difference between both compressions in any of the situations, subjects showed a preference for dual compression.

Moore., (2008) suggested some conclusions regarding the choice of compressions speed in hearing aids:

- The benefit from compression was greatest among individuals, who experience a wide range of sound levels within short periods of time,
- Slow compression generally leads to higher listening comfort than fast compression,
- The benefit from fast compression varies across individuals, and those with high cognitive ability are able to benefit from fast compression to take advantage of temporal dips in a background sound. It is argued that listening in the dips depends on the ability to process the temporal fine structure of sounds. It is proposed that a test of the ability to process temporal fine structure might be useful for selecting compression speed for an individual.

Higgins, Searchfield, and Coad (2012) studied the effect of fast and slow acting WDRC and adaptive dynamic range optimization (ADRO) speech in noise perception. Two receiver-in-the-ear hearing aids were compared: one using 32-channel adaptive dynamic range optimization (ADRO) and the other wide dynamic range compression (WDRC) with dual fast (4 channel) and slow (15 channel) processing. The manufacturer's first-fit settings based on subject's audiograms were used in both cases. Results were obtained from 18 participants on a quick speech-in-noise (QuickSIN) task. ADRO processing speech recognition in noise was better compared to the multichannel WDRC processing. The suggested reason for the better performance of the ADRO hearing aids was lesser fluctuation in output when there was a change in sound dynamics.

Henceforth from review of literature it is evident that there are studies on temporal aspects of speech, and its importance in perception. There are also evidences on perception of speech through different hearing aid compression. But there is dearth of literature focusing on the combined effect of temporal cues and hearing aid compressions together, hence this study focuses regarding these aspects.

Chapter 3

Method

The present study aimed to determine the performance on the identification of hearing aid processed speech with the use temporal fine structure cues and temporal envelope cues in auditory chimeric Malayalam sentences by the normal hearing individual.

3.1 Participants

The number of participants proposed for the study was thirty, however, thirty-five (15 male and 20 females) normal hearing individuals who were native speakers of Malayalam within the age range of 18-30 years (Mean = 22 years, SD = ±2.58) participated in the study**.**

3.1.1 Inclusion criteria

- Hearing threshold: Air conduction hearing thresholds should be within 15 dB at octave frequencies between 250Hz to 8000Hz
- Speech identification scores proportional to the hearing threshold
- Immittance audiometry: ‘A’ type tympanogram in the presence of acoustic reflexes in both ears at 500 Hz, 1000Hz, 2000Hz and 4000Hz
- Uncomfortable level of speech greater than or equal to 100 dBHL in both ears
- Presence of transient Otoacoustic Emissions in both ears
- No history or presence of any external or middle ear problem
- No history or presence of any neurological problem
- Native speaker of Malayalam

Prior written consent forms are taken from the participants for their willingness to participate in the study.

3.2 Equipment

The following equipments were used in the study

- A calibrated two channel diagnostic audiometer Grason –Stadler model GSI-61 coupled with acoustically matched TDH 39 headphones housed in MX-41/AR and Radio ear B-71 bone vibrator were used to estimate pure-tone threshold, speech recognition thresholds, speech identification thresholds and uncomfortable levels.
- Calibrated middle ear analyzer GSI tympstar version 2 was used for tympanometry and reflexometry.
- Otoacoustic measurements were carried out using ILOV 6 equipment.
- A digital BTE hearing aid with accessibility to use 2 programs was used.
- KEMAR (Knowles Electroacoustic Manikin for Auditory Research) was used for recording the processed speech from the hearing aid.
- Bruel & Kjær (BZ-5503) Sound level meter was used for checking the presentation level of the chimeric sentence for recording the processed stimuli and also to pick up the output of the hearing aid.

A laptop, loaded with the following softwares:

- Hilbert transform using MATLAB software [MATLAB 7.12.0 (R2011a)]
- Cubase (version 2.0.2) was used for the presentation of the chimeric sentence
- Bruel & Kjær (BZ-5503) software was used to record the processed stimulus.
- Adobe Audition (version 3.0) was used for the presentation of the stimulus

- PRAAT (version 6.0.39) software was used for recording the response of the subject
- A personal computer installed with NOAH 3.0 was used for programming the hearing aids.

3.3 Environment

All audiological tests are carried out in a sound-treated double room where the noise levels are within permissible limits (ANSI, 1999).

3.4 Material

Sentences for preparing chimeric list was selected from “Sentence list in Malayalam” (Sreeraj & Kishore, 2014). The sentence list consists of 12 full lists of 120 sentences all together with 10 sentences in each list. The sentences were selected such that the total number of words in each sentence was limited to four in a sentence not more than that. Total of eighty pairs of sentences was taken to prepare a speech – speech chimera across eight frequency bands which includes one, four, six, eight, sixteen, twenty-eight, thirty-two, and sixty-four.

3.5 Procedure

The study is carried out in 4 phases.

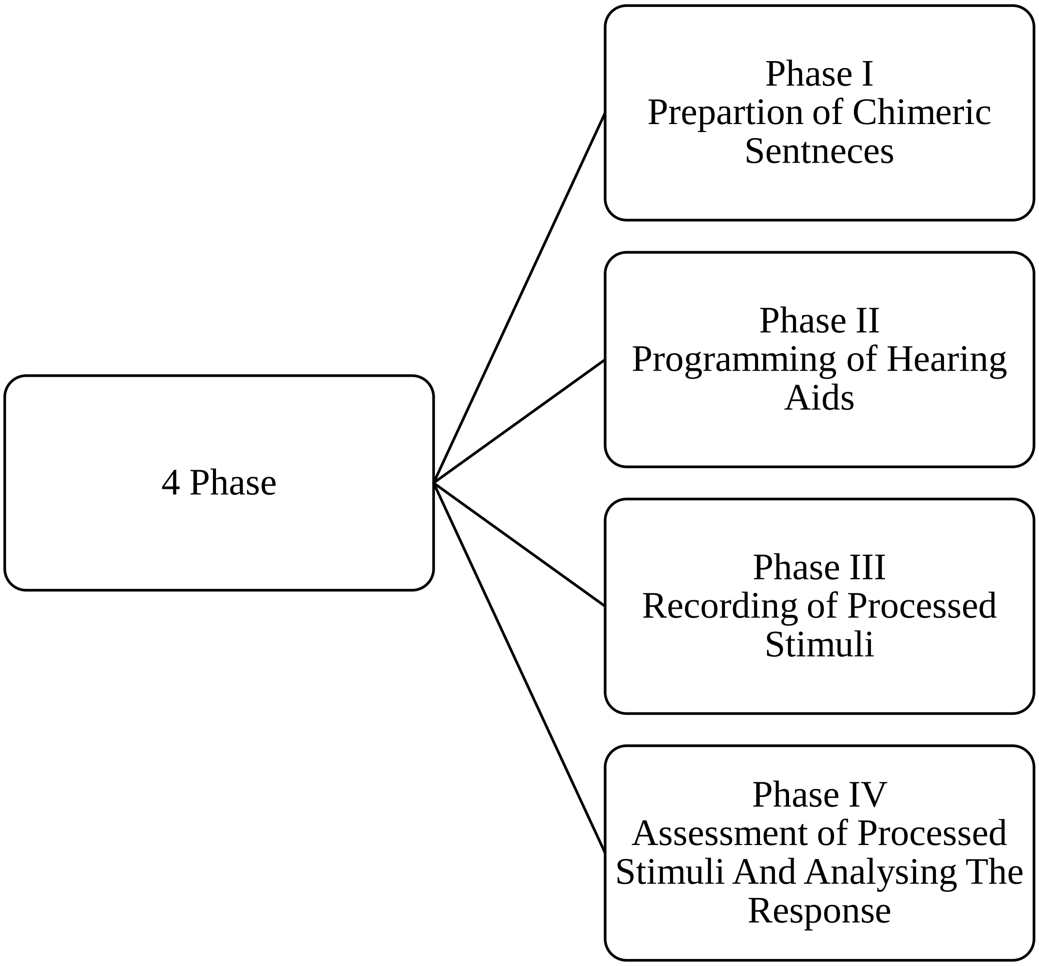

Figure 3.1 Flow chart depicting different phases of the study

3.5.1 Phase 1: Preparation of speech - speech chimera sentences in Malayalam

- The sentence list from the material was grouped for the envelope and fine structure extraction for the preparation of speech to speech chimeras.
- The selected eighty pairs of sentences were processed using Hilbert transform to extract the temporal cues such as envelope and fine structure. Hilbert transform was mainly used to derive envelope function or instantaneous amplitude of a signal. It mainly represents a filter without affecting the gain (Yost & Fay, 2007).

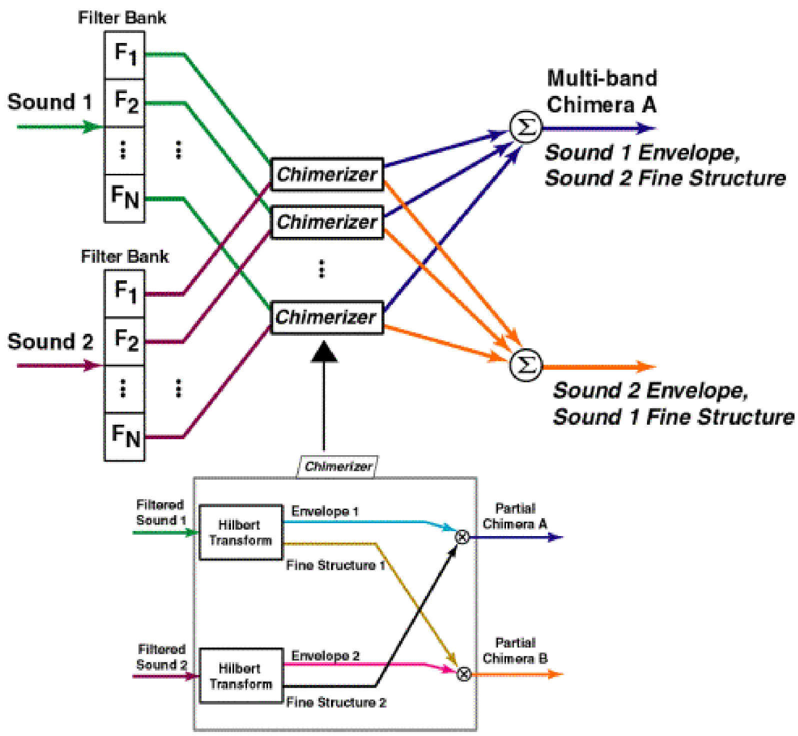

Figure 3.2. Diagrammatic representation of the preparation of chimeric stimuli

Note: Reprinted from Perception of auditory chimeras by Smith, Oxenham and Deglutte, 2001, retrieved from http://research.meei.harvard.edu/chimera/More.html, Copyright 2001 by Association for Research in Otolaryngology

The figure 3.2 shows the process behind the computation of Hilbert transform. The steps involved are as follows:

- First, the system calculates the Fourier transform of the given signal x(t).
- Second, it rejects the negative frequencies.
- Finally, it calculates the inverse Fourier transform, and the result will be a complex-valued signal where the real and the imaginary parts form a Hilbert-transform pair.

For example: When x(t) is narrow-banded, |z(t)| can be regarded as a slow- varying envelope of x(t) while the phase derivative $\partial t\ [\tan-1(y/x)]$ is an instantaneous frequency. Thus, the Hilbert transform can be interpreted as a way to symbolize a narrow-band signal in terms of amplitude and frequency modulation (Shi, Lee, Liu, Yang, & Wang, 2011). After obtaining envelope and fine structure for each sentence, these temporal cues were exchanged with each other in order to make speech-speech auditory chimeric sentences. For example, the envelope of sentence one is combined with the fine structure of sentences two to make one chimeric sentence. Likewise, cues are exchanged between all sentences and 80 chimeric sentences are made. The same procedure is followed to make speech-speech chimera for words.

- Each chimeric sentence was synthesized with 1, 4, 6, 12, 16, 28, 32 and 64 frequency bands and a total of 80 chimeric sentences were created.
- The total number of syllables in each sentence is limited to 8 to 9 syllables and each word in a sentence will not have more than 3-4 syllables.
- Due to the reduction in the number of pair of sentence list, 4 sentence list were randomized to prepare the 32 and 64 band chimeric sentence list.

3.5.2 Phase II: Programming of the hearing aid

- Programming of hearing aid was done using a Noah 3.0 software where the hearing aid was coupled to the personal computer through Hi-pro and the programming was simulated to a 40 dBHL loss of flat configuration because to retain the compression characteristics across the frequencies.

- The digital BTE aid was programmed with 2 programs, the syllabic compression and dual compression which was designated as program 1 and 2 respectively with a prescriptive procedure of NAL-NL 1.

3.5.3 Phase III: Preparation of hearing aid processed stimuli.

- The prepared chimeric sentences were presented through the speaker routed by audiometer to the hearing aid and the output is recorded. The setup includes a microphone to pick the output of the hearing aid in the KEMAR which is coupled with the hearing aid, with the help of 2 cc coupler in the KEMAR a connection was made to the microphone of the SLM, where the level was maintained and also the recording is carried out. The stimuli are presented through the loudspeaker at the 0° angle through Cubase software (version 2.0.2) system.

- During the process of recording the hearing aid was set to program 1 and program 2 separately and the list of sentences was recorded with both the programs separately.

3.5.4 Phase IV: Administration of processed sentence list

- All stimuli were presented binaurally via headphones at a comfortable listening level of 65dBHL. Practice sessions were provided using 3 chimeric sentences
- Each participant was provided with 160 chimeric sentences which were processed in two programs respectively. The participant's task is to repeat the sentence that was perceived. The responses were recorded.

- The recorded sentences were subjected to further analysis in the presence of a native Malayalam speaker were speech identification scores were measured for each sentence.
- Based on the number of keywords repeated the score was given from 0 to 4 (0 being no words repeated and 4 being all the 4 words repeated correctly)
- After scoring these scores were converted to a percentage to estimate the speech identification scores.

Chapter 4

Results

The aim of the present study was to investigate if there are any effect of syllabic compression and dual compression on nonlinear hearing aid processed Malayalam chimeric sentences in normal hearing individuals. The data obtained from 35 participants were subjected to statistical analysis for the perception of TFS and ENV cues between two compression systems processed chimeric sentences across 8 frequency bands i.e., 1, 4, 6, 8, 16, 24, 32, & 64. All the collected data were tabulated and statistically analyzed in the SPSS platform version 21.

All the collected data were first subjected to descriptive statistical analysis where the overall mean, median, and standard deviation were calculated across frequency band between dual and syllabic compression processed chimeric sentence list. The data were tabulated and mentioned in Table 4.1 below.

Table 4.1 Overall Raw Mean, Median and Standard Deviation of Chimeric sentences across different frequency bands and within compression.

Stimulus	Mean (%)	SD	Median (%)
Du 1	7.00	4.955	5.00
Du 4	13.07	7.07	15.00
Du 6	10.35	5.85	10.00
Du 8	17.85	8.47	17.50
Du 16	78.64	10.83	77.50
Du 24	88.92	4.89	90.00
Du 32	95.07	3.76	95.00
Du 64	99.71	0.80	100.00
Sy 1	7.14	4.62	7.50
Sy 4	14.85	6.99	15.00
Sy 6	11.28	5.15	10.00
Sy 8	18.57	8.42	17.50
Sy 16	80.00	9.17	80.00
Sy 24	89.42	6.03	90.00
Sy 32	95.50	4.32	95.00
Sy 64	99.71	0.80	100.00

Note: Du Du= Dual compression, Sy = Syllabic compression, 1-64 indicates a number of frequency bands.

The data table 4.1 depicts the overall mean, standard deviation and median score across frequency band between dual and syllabic compression. As it is shown in the table 4.1the SD across all the parameters were so high, hence the median values were employed for further statistical analysis. The graphical representation of the median percentage scores was shown in figure 4.1 below.

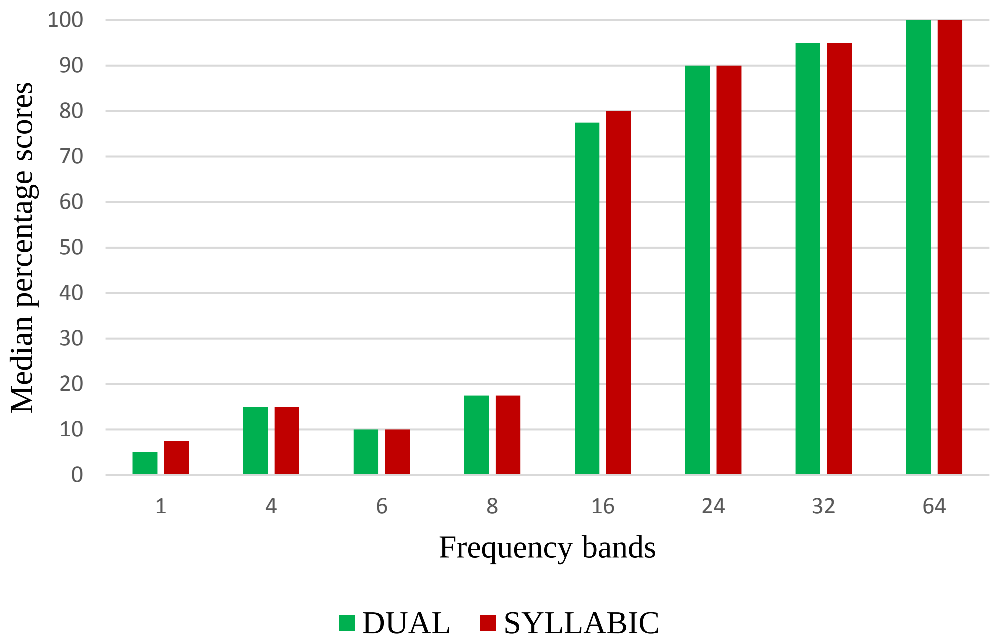

Figure 4.1. Graphical representation of Median scores of Chimeric sentences across different frequency bands between dual and syllabic compression

Based on the data table 4.1, the statistical analysis was carried out to assess the following aspects:

- The frequency of identification of temporal fine structure and envelope cues across the frequency bands between the dual and syllabic compression system processed chimeric sentences.
- Effect of dual and syllabic compression on envelope cue identification across the frequency bands in chimeric sentences.
- Effect of dual and syllabic compression on temporal fine structure cue identification across the frequency bands in chimeric sentences.
- Comparison of dual and syllabic compression on the identification of Malayalam chimeric sentence

4.1 Identification of temporal fine structure and envelope cues across the frequency bands between the dual and syllabic compression system processed chimeric sentences.

The data were calculated for the frequency of participants correctly identified any-one cues were measured across conditions and it is tabulated in table 4.2

Table 4.2 Details of participant frequency (%) for identification of fine structure and envelope cues in dual and syllabic compression system across frequency bands for chimeric sentences.

Frequency band	Dual		**Syllabic**	
	Fine Structure (%)	**Envelope (%)**	**Fine Structure (%)**	**Envelope (%)**
S 1	94.28	5.72	94.28	5.72
S 4	2.86	97.14	0	100
S 6	0	100	0	100
S 8	0	100	0	100
S 16	0	100	0	100
S 24	0	100	0	100
S 32	0	100	0	100
S 64	0	100	0	100

Note: S = Chimeric sentence list, 1- 64 indicates the number of frequency bands

After the analysis of data on participants frequency of response. The cues are segregated and from table 4.2 it was inferred that the at low or reduced frequency bands the TFS perception was predominant however as the number frequency bands gradually increased the perception of envelope cues were more or high and hence it was evident from table 4.2 that the TFS cue exhibited no response more compared to the ENV cues hence the ENV cues at all bands and the 1 band TFS cue are subjected for further inferential statistical analysis of the data.

Before the inferential statistical analysis, all the data were subjected to normality test for further selection of inferential statistics, Shapiro-Wilk's test of normality was done along with outlier analysis for the significance of $p > 0.05$. But the data was not followed a normal distribution and also outliers were not seen so the non-parametric test was selected for further statistical analysis of data.

4.2 Effect of dual and syllabic compression on envelope cue identification across the frequency bands in chimeric sentences.

As the frequency of response across the envelope cue was found to be more predominant as in the data table 4.2. The effect of the dual and syllabic compression on envelope cue across the frequency bands was subjected to analysis. The non-parametric test, Freidman's test was carried out between the dual and syllabic compression for a significance of $p < 0.01$ as depicted in table 4.3.

Table 4.3 Result of Friedman test for ***c****omparison across frequency bands in envelope cues between dual and syllabic compression on chimeric sentence identification.*

Condition	χ^2
Dual compression	231.140*
Syllabic compression	233.310*

*Note: * indicates p < 0.01*

The data analysis showed a significant difference across frequency band in both the dual and syllabic compression condition on the envelope cue identification. Hence, the pairwise comparison of both the dual and syllabic compression processed chimeric sentence identification data across frequency band was subjected to statistical analysis using Wilcoxon's signed rank test with a significance of $p < 0.01$ which is depicted in table 4.4.

Table 4.4 Result of Wilcoxon signed rank test for ***c****omparison across frequency bands within dual compression in envelope cue on chimeric sentence identification*

	1 Du	**4 Du**	**6 Du**	**8 Du**	**16 Du**	**24 Du**	**32 Du**	**64 Du**
1 Du		5.033*	5.173*	5.167*	5.166*	5.184*	5.195*	5.555*
4 Du			2.071*	2.687*	5.165*	5.168*	5.168*	5.175*
6 Du				3.873*	5.171*	5.172*	5.170*	5.180*
8 Du					5.165*	5.166*	5.173*	5.166*
16 Du						4.429*	4.944*	5.093*
24 Du							4.626*	5.118*
32 Du								4.582*
64 Du								

*Note: * indicates $p < 0.01$, Du = dual compression*

From the pairwise comparison of dual compression processed chimeric sentence identification across frequency band showed a significant difference between each frequency band on the perception of envelope cue.

The syllabic compression data also showed a significant difference in the perception of envelope cues on the Friedman's test which was depicted in table 4.3. The pairwise comparison across frequency bands for the envelope cue identification on syllabic compression processed chimeric sentences were subjected to the statistical test of

Wilcoxon's signed ranked test with a significance of $p < 0.01$ which is depicted in table 4.5.

Table 4.5 Result of Wilcoxon signed rank test for ***c****omparison across frequency bands within dual compression in envelope cue on chimeric sentence identification*

	1 Sy	**4 Sy**	**6 Sy**	**8 Sy**	**16 Sy**	**24 Sy**	**32 Sy**	**64 Sy**
1 Sy		5.172*	5.103*	5.173*	5.171*	5.181*	5.212*	5.600*
4 Sy			2.132*	1.964*	5.167*	5.171*	5.175*	5.169*
6 Sy				3.873*	5.171*	5.172*	5.170*	5.180*
8 Sy					5.164*	5.168*	5.166*	5.175*
16 Sy						4.353*	5.175*	5.176*
24 Sy							4.210*	5.047*
32 Sy								4.530*
64 Sy								

*Note: * indicates p <0.01, Sy = Syllabic compression, 1 to 64 bands*

In table 4.5 also the same trend follows similarly to the dual compression processed chimeric sentences that the pairwise comparison showed a significant difference across frequency band on the identification of envelope cue on syllabic compression processed chimeric sentences as depicted in table 4.2.

4.3 Effect of dual and syllabic compression on temporal fine structure cue identification within the frequency bands in chimeric sentences.

As the data from the Table 4.2 depicted that the temporal fine structure cues were predominant only in 1 band and hence it was subjected to statistical analysis within the frequency band but due to the lack of comparison data within TFS on both the dual and syllabic compression processed chimeric sentences the 1 band data on the envelope cue from both dual and syllabic compression processed chimeric sentence were taken for comparison to rule out the predominant cue on this band. Hence non-parametric test Wilcoxon's signed rank test was carried out within the band across the condition at a significance $p < 0.01$ which is depicted in table 4.6.

Table 4.6 Result of Wilcoxon signed rank test for ***c****omparison within a frequency band in temporal fine structure cues within dual and within syllabic compression on chimeric sentence identification.*

1 band	**│Z│**
Du TFS – Du ENV	4.881*
Sy TFS – Sy ENV	4.890*

*Note: * indicates p <0.01, Du = dual compression, Sy = syllabic compression, TFS = Temporal Fine Structure and ENV = Envelope*

The results showed a significant difference between the temporal fine structure and envelope cue identification across the dual and syllabic compression processed chimeric sentences list.

4.4 Comparison of dual versus syllabic compression on the identification of Malayalam chimeric sentence

The statistical results on within-group comparison among dual and syllabic condition on the identification of TFS and ENV cues across frequency bands revealed a significant difference. Hence, between-group comparisons were carried out among TFS and ENV cues using inferential statistics. From the table 4.2, it showed that the frequency of response at 1 frequency band was more for TFS cue identification. However, the ENV cues are prevalent across all the frequency and hence the 1 frequency band TFS cue was subjected to statistical analysis first between the dual and syllabic compression condition. The non-parametric test Wilcoxon's signed ranked was carried out at a significance of $p < 0.01$ & $p < 0.05$ and the data is depicted in table 4.7

Table 4.7 Result of Wilcoxon signed rank test for ***c****omparison between dual and syllabic compression in temporal fine structure cue within 1 frequency band on chimeric sentence identification.*

1 band	**│Z│**
Du TFS – Sy TFS	0.027

Note: Du = dual compression, Sy = syllabic compression, and TFS = Temporal Fine Structure

The results of Wilcoxon signed rank test as depicted in the table 4.7 showed no significant difference between dual and syllabic compression processed chimeric sentence identification. However, from the table 4.1 on overall median score showed a marginal difference between both the dual and syllabic compression system, being the syllabic compression processed chimeric sentence list exhibited better score.

Secondly, the entire frequency band from the ENV cues were subjected for pairwise comparison of data both within the band and between the band using the Wilcoxon signed rank test with a significance of $p < 0.01$ & $p < 0.05$ which was depicted in the data Table 4.8

*Table 4.8 Result of Wilcoxon signed rank test for **c**omparison between dual and syllabic compression in envelope cue within and between frequency bands on chimeric sentence identification.*

	1 Du	4 Du	6 Du	8 Du	16 Du	24 Du	32 Du	64 Du
1 Sy		5.031*	5.173*	5.169*	5.166*	5.185*	5.191*	5.555*
4 Sy		0.923	2.598**	1.882**	5.166*	5.166*	5.166*	5.174*
6 Sy			0.851	3.567*	5.164*	5.169*	5.172*	5.173*
8 Sy				0.692	5.165*	5.166*	5.166*	5.170*
16 Sy					1.045	4.681*	5.099*	5.177*
24 Sy						0.987	4.155*	5.044*
32 Sy							0.643	4.537*
64 Sy								

*Note: * indicates p < 0.01, **indicates & p < 0.05, Du = Dual compression, Sy = Syllabic compression*

The result of Wilcoxon signed rank test within the frequency band showed no significant difference within the frequency band but from the data Table 4.1 on overall median data showed some amount of marginal difference being the syllabic compression data better with the between frequency band showing a significant difference between the dual and syllabic compression processed chimeric sentence list being syllabic compression processed chimeric sentence list exhibiting better perception of Malayalam chimeric sentences.

Chapter 5

Discussion

In general, any speech stimuli contain two cues in it which are factored separately based on their slow varying nature called the envelope or ENV cues and the fast or rapidly changing temporal fine structure or TFS cues (Moore & Sęk, 2009; Rosen, 1992). To separately assess the importance of these cues basically two algorithms were employed with one being the peak clipping and filtering (Licklider & Pollack, 1948). In this technique, the peaks were clipped and the signals was subjected to low pass or high pass filtering and then the signals are made to vocode with a tone or noise. The major disadvantage of this method as believed by the researchers is that it provides the differentiation between the cues for perception however it is not, instead it creates more of distortion. Hence, another technique came in existence for extraction of different cues called the Hilbert transform (Bracewell & Bracewell, 1986) where it uses the mathematical calculation and FFT to extract the two features from a speech stimuli separately and a speech to speech chimaeras are prepared across frequency bands. However, studies as reported in normal hearing individual across different languages Malayalam language (Indu & Devi, 2016) on English language (Smith et al., 2002) and mandarin Chinese language (Xu et al., 2005) revealed that the predominance is more for TFS cue rather than the ENV cue. Naveen and Devi, (2017) on the Kannada language with the earlier study as a base on ruling out the relative importance across cues found ENV to be perceived better with perception predominance start from 4 frequency bands which is similar to that of study on the English language.

However, there are no study which addressed how the cues of ENV and TFS this cue help in case of hearing-impaired individual perception of speech when perceived through two different (fast and slow acting) compression system. The difference between the compression systems are basically based on the attack and release time of the hearing aid fast acting or syllabic compression system had increased attack and release time, the slow acting or dual compression system had less attack and release time. Geetha and Manjula, (2005) studied the effect of dual and syllabic compression system on speech identification score. Also, Merin and Devi, (2015) employed the dual and syllabic compression system processed music stimuli to assess the music perception ability on the normal hearing individual.

Hence, considering all the above-mentioned research work as a base this study was carried out to assess the influence of this two-compression system on ENV and TFS cue identification using the hearing aid processed speech stimuli on the normal hearing individual. The results of this study focused firstly on relative importance between the perception of TFS vs ENV cues revealed that the perception of ENV cues is predominant after 4 frequency bands and TFS cues are predominant at 1 frequency band between both the compression system and also across the frequency band there is a significant improvement in the perception of ENV cues from 4 through 64 frequency bands, these findings are supported and followed the same trend of study on English language (Smith et al., 2002) and Kannada language (Naveen & Devi, 2017). However, the finding are contrasted with the earlier study on Malayalam language (Indu & Devi, 2016) one of the possible reason is that in the earlier study the vocoder technique was used for extracting information. This present study used Hilbert transform to process for feature extraction follows different algorithm which might be the possible reason for variation in the results and also possibly due to the envelope reconstruction which starts from the 4 frequency bands itself. (Gilbert et al., 2006)

The second objective of this study was to rule out if there was any difference between the two compression systems on speech identification. The results of this study are in support with Geetha and Manjula, (2015) that there was no significant difference between both the compression processed speech stimuli on identification of chimeric sentence list. But this study results were contradicting to the statement on the preference towards dual compression system by the hearing-impaired individual as reported by Geetha and Manjula, (2015) and Merin and Devi, (2015) were dual compression system sounded to be better on the basis of loudness parameters for perception of processed music. The present study showed some marginal improvement in syllabic compression compared to dual compression through the overall descriptive statistical value. However, Neuman et al., (1995) has reported syllabic compression system to be better.

Chapter 6

Summary and Conclusion

To summarize the study any speech stimuli, consist of two features in it and both the features had its own predominance across frequency bands hence to study these two algorithms are usually present that is the peak clipping with filtering and the Hilbert transform. In the present study, the Hilbert transform is used for the extraction and mixing of the features from the Malayalam sentence list is used to prepare a speech to speech chimeric sentence list. The current study, was carried out using 4 channel hearing aid with a dual and syllabic compression option available in it. The hearing aid was simulated to a 40dB flat loss with 2 individual programs being dual and syllabic compression independent to each other. In order to prepare the processed stimuli for the study the hearing aid was provided with the chimeric sentences through the loudspeaker at an angle of 0° and the output was recorded with the help of KEMAR and saved as a wave file through adobe audition. The data are collected on the normal hearing individual with circumaural headphones coupled through the audiometer and the stimuli was presented at 40 dB SL to the participants and the response of the individual was recorded through PRAAT software and saved as.WAV file for further analysis.

Each of the sentences contains 4 words in it, and all the words are keywords from the sentences. Each list consists of 10 sentences in it and the scoring was done with the help of native Malayalam speaker. A score of 1 was given for each keyword responded. So in total each sentence was scored for 4 with each sentence list contain a score of overall 40. After scoring all the scores were converted to 100% to get the speech identification score and the data were tabulated and statistically analyzed using the SPSS software version 21.

In the statistical analysis the data were subjected to both the descriptive and inferential statistics, and in the descriptive statistics the overall mean, standard deviation and median are computed along with that a normality and outlier analysis are also done. Based on the normality check non-parametric test are subjected to all the data and it revealed a significant difference across the frequency band in perception with the ENV cues being perceived better with an increase in the frequency band with TFS cue predominant with the low-frequency band. And between the compressions, no significant difference was observed but marginal difference was observed between the conditions with the better being the syllabic compression.

To conclude the ENV reconstruction leads to the better perception of ENV cues from the 4 band itself through 64 bands which actually not persistent at the low frequency band. It also supports the earlier research findings that the trend might not be the same in case of individuals with hearing impairment increased frequency bands are required for better perception of speech. However, the tradeoff between the two compressions might also vary from the normal hearing individual to hearing impaired individual.

Clinical implication:

Based on the present study it is noted that the ENV cues are perceived better because of the reconstruction happening at the level of BM and also for a better perception of speech. The frequency bands should be more to overcome the limited phase locking ability till 4 kHz which impairs TFS perception above it. And also while programming hearing aid in case of hearing-impaired individual syllabic compression can be used for better perception of speech but the efficacy needs to be assessed by doing furthermore research on hearing impaired individual.

Future direction:

- The study is mainly carried out on normal hearing individuals with simulated processed stimuli, this could be extended to the hearing-impaired subjects within and across different degrees of loss.
- The compression ratio between the compression systems can be varied.
- There is supporting study for syllabic compression being better in noisy situation hence further research work can be done across different noisy conditions.

Reference

Benson, D., Clark, T. M., & Johnson, J. S. (1992). Patient experiences with multiband full dynamic range compression. *Ear and Hearing*, *13*(5), 320–30. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1487092

Bracewell, R., & Bracewell, R. (1986). *The Fourier transform and its applications*. (McGraw-Hill Book Company, Ed.). Newyork, USA. Retrieved from http://www.academia.edu/download/44001876/34957138.pdf

Buss, E., Hall, J. W., & Grose, J. H. (2004). Temporal Fine-Structure Cues to Speech and Pure Tone Modulation in Observers with Sensorineural Hearing Loss. *Ear and Hearing*, *25*(3), 242–250. https://doi.org/10.1097/01.AUD.0000130796.73809.09

Buss, E., Iii, J. W. H., & Grose, J. H. (2004). Temporal Fine-Structure Cues to Speech and Pure Tone Modulation in Observers with Sensorineural Hearing Loss, 242–250. https://doi.org/10.1097/01.AUD.0000130796.73809.09

Byrne, D., Dillon, H., Ching, T., Katsch, R., & Keidser, G. (2001). NAL-NL1 procedure for fitting nonlinear hearing aids: characteristics and comparisons with other procedures. *Journal of the American Academy of Audiology*, *12*(1), 37–51.

Chen, B. K., Clark, G. M., & Jones, R. (2003). Evaluation of trajectories and contact pressures for the straight nucleus cochlear implant electrode array - a two-dimensional application of finite element analysis. *Medical Engineering & Physics*, *25*(2), 141–7. https://doi.org/10.1016/S1350-4533(02)00150-9

Dallos, P., Popper, A., Fay, R., & Dallos, P. (1996). The Cochlea: Springer Handbook of Auditory Research. Springer-Verlag. Retrieved from https://www.scholars.northwestern.edu/en/publications/the-cochlea-springer-handbook-of-auditory-research

Dreschler, W. A. (1989). Phoneme Perception via Hearing Aids with and without Compression and the Role of Temporal Resolution. *International Journal of Audiology*, *28*(1), 49–60. https://doi.org/10.3109/00206098909081610

Eisenberg, L. S., Shannon, R. V., Martinez, A. S., Wygonski, J., & Boothroyd, A. (2000). Speech recognition with reduced spectral cues as a function of age. *The Journal of the Acoustical Society of America*, *107*(5), 2704. https://doi.org/10.1121/1.428656

Flanagan, J. L. (1980). Parametric coding of speech spectra. *The Journal of the Acoustical Society of America*, *68*(2), 412–419. https://doi.org/10.1121/1.384752

Florentine, M., Buus, S., Scharf, B., & Zwicker, E. (1980). Frequency Selectivity in Normally-Hearing and Hearing-Impaired Observers. *Journal of Speech Language and Hearing Research*, *23*(3), 646. https://doi.org/10.1044/jshr.2303.646

Fu, Q., Zeng, F., Shannon, R. V, Soli, S. D., & Shannon, R. V. (2006). Importance of tonal envelope cues in Chinese speech recognition, *505*(1998).

Ghitza, O. (2001). On the upper cutoff frequency of the auditory critical-band envelope detectors in the context of speech perception. *The Journal of the Acoustical Society of America*, *110*(3), 1628–1640. https://doi.org/10.1121/1.1396325

Gilbert, G., Lorenzi, C., Gilbert, G., & Lorenzi, C. (2006). The ability of listeners to use recovered envelope cues from speech fine structure The ability of listeners to use recovered envelope cues from speech fine structure, *2438*. https:// doi.org/ 10.1121 /1.2173522

Giraud, A.-L., Lorenzi, C., Ashburner, J., Wable, J., Johnsrude, I., Frackowiak, R., & Kleinschmidt, A. (2000). Representation of the Temporal Envelope of Sounds in the Human Brain. *Journal of Neurophysiology*, *84*(3), 1588–1598. https://doi.org / 10.1152 /jn.2000.84.3.1588

Gnansia, D., Péan, V., Meyer, B., & Lorenzi, C. (2009). Effects of spectral smearing and temporal fine structure degradation on speech masking release. *The Journal of the Acoustical Society of America*, *125*(6), 4023–4033. https://doi.org /10.1121/ 1.3126344

Heinz, M. G., & Swaminathan, J. (2009). Quantifying Envelope and Fine-Structure Coding in Auditory Nerve Responses to Chimaeric Speech. *Journal of the Association for Research in Otolaryngology*, *10*(3), 407–423. https://doi.org /10.1007/s10162-009-0169-8

Hickson, L. M. H. (1994). Compression Amplification in Hearing Aids. *American Journal of Audiology*, *3*(3), 51. https://doi.org/10.1044/1059-0889.0303.51

Higgins, P., Searchfield, G., & Coad, G. (2012). A Comparison Between the First-Fit Settings of Two Multichannel Digital Signal-Processing Strategies: Music Quality Ratings and Speech-in-Noise Scores. *American Journal of Audiology*, *21*(1), 13. https://doi.org/10.1044/1059-0889(2011/10-0034)

Hnath-Chisolm, T. E., Laipply, E., & Boothroyd, A. (1998). Age-Related Changes on a Children's Test of Sensory-Level Speech Perception Capacity. *Journal of Speech Language and Hearing Research*, *41*(1), 94. https://doi.org/10.1044/jslhr.4101.94

Hopkins, K., & Moore, B. C. J. (2007). Moderate cochlear hearing loss leads to a reduced ability to use temporal fine structure information. *The Journal of the Acoustical Society of America*, *122*(2), 1055–1068. https://doi.org/10.1121/1.2749457

Hopkins, K., & Moore, B. C. J. (2009). The contribution of temporal fine structure to the intelligibility of speech in steady and modulated noise. *The Journal of the Acoustical Society of America*, *125*(1), 442–446. https://doi.org/10.1121/1.3037233

Hopkins, K., Moore, B. C. J., & Stone, M. A. (2008). Effects of moderate cochlear hearing loss on the ability to benefit from temporal fine structure information in speech. *The Journal of the Acoustical Society of America*, *123*(2), 1140–1153. https://doi.org/10.1121/1.2824018

Joris, P. X., & Yin, T. C. T. (1992). Responses to amplitude- modulated tones in the auditory nerve of the cat. *The Journal of the Acoustical Society of America*, *91*(1), 215–232. https://doi.org/10.1121/1.402757

King, A. B., & Martin, M. C. (1984). Is AGC beneficial in hearing aids? *British Journal of Audiology*, *18*(1), 31–38. https://doi.org/10.3109/03005368409078926

Licklider, J. C. R., & Pollack, I. (1948). Effects of Differentiation, Integration, and Infinite Peak Clipping upon the Intelligibility of Speech. *The Journal of the Acoustical Society of America*, *20*(1), 42–51. https://doi.org/10.1121/1.1906346

Loizoumichael, P. C., Tu, D., & Dorman, M. (2006). On the number of channels needed to understand speech, *2097*(1999). https://doi.org/10.1121/1.427954

Lorenzi, C., Debruille, M., Garnier, S., Fleuriot, P., Moore, M. C. J., & Moore, B. C. J. (2010). Abnormal processing of temporal fine structure in speech for frequencies where absolute thresholds are normal Abnormal processing of temporal fine structure in speech for frequencies where absolute thresholds are normal (L), *27*(2009). https://doi.org/10.1121/1.2939125

Lorenzi, C., Dumont, A., & Fu¨llgrabe, C. (2000). Use of Temporal Envelope Cues by Children With Developmental Dyslexia. *Journal of Speech Language and Hearing Research*, *43*(6), 1367. https://doi.org/10.1044/jslhr.4306.1367

Lorenzi, C., Gilbert, G., Carn, H., Garnier, S., & Moore, B. C. J. (2006). Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(49), 18866–9. https://doi.org/10.1073/pnas.0607364103

Moore, B. C. J. (2008). The Choice of Compression Speed in Hearing Aids: Theoretical and Practical Considerations and the Role of Individual Differences. *Trends in Amplification*, *12*(2), 103–112. https://doi.org/10.1177/1084713808317819

Moore, B. C. J., & Glasberg, B. R. (1987). Formulae describing frequency selectivity as a function of frequency and level, and their use in calculating excitation patterns. *Hearing Research*, *28*(2–3), 209–225. https://doi.org/10.1016/0378-5955(87)90050-5

Moore, B. C. J., Glasberg, B. R., & Hopkins, K. (2006). Frequency discrimination of complex tones by hearing-impaired subjects: Evidence for loss of ability to use temporal fine structure. *Hearing Research*, *222*(1–2), 16–27. https:// doi.org /10.1016/J.HEARES.2006.08.007

Moore, B. C. J., Glasberg, B. R., & Stone, M. A. (1999). Use of a loudness model for hearing aid fitting: III. A general method for deriving initial fittings for hearing aids with multi-channel compression. *British Journal of Audiology*, *33*(4), 241–258. https://doi.org/10.3109/03005369909090105

Moore, B. C. J., & Lorenzi, C. (2006). Speech perception problems of the hearing impaired reflect inability to use temporal fine structure, 5–8.

Moore, B. C. J., Peters, R. W., & Stone, M. A. (1999). Benefits of linear amplification and multichannel compression for speech comprehension in backgrounds with spectral and temporal dips. *The Journal of the Acoustical Society of America*, *105*(1), 400. https://doi.org/10.1121/1.424571

Moore, B. C. J., & Sęk, A. (2009). Sensitivity of the human auditory system to temporal fine structure at high frequencies. *The Journal of the Acoustical Society of America*, *125*(5), 3186. https://doi.org/10.1121/1.3106525

Nazzi, T., Iakimova, G., Bertoncini, J., Frédonie, S., & Alcantara, C. (2006). Early segmentation of fluent speech by infants acquiring French: Emerging evidence for crosslinguistic differences. *Journal of Memory and Language*, *54*(3), 283–299. https://doi.org/10.1016/J.JML.2005.10.004

Nazzi, T., Jusczyk, P. W., & Johnson, E. K. (2000). Language Discrimination by English-Learning 5-Month-Olds: Effects of Rhythm and Familiarity. *Journal of Memory and Language*, *43*(1), 1–19. https://doi.org/10.1006/JMLA.2000.2698

Neuman, A. C., Bakke, M. H., Mackersie, C., Hellman, S., & Levitt, H. (1995). Effect of release time in compression hearing aids: Paired- comparison judgments of quality. *The Journal of the Acoustical Society of America*, *98*(6), 3182–3187. https://doi.org/10.1121/1.413807

Nie, K., Barco, A., & Zeng, F.-G. (2006). Spectral and Temporal Cues in Cochlear Implant Speech Perception. *Ear and Hearing*, *27*(2), 208–217. https://doi.org/10.1097/01.aud.0000202312.31837.25

Nuttall, A. L., Christian Brown, M., Masta, R. I., & Lawrence, M. (1981). Inner hair cell responses to the velocity of basilar membrane motion in the guinea pig. *Brain Research*, *211*(1), 171–174. https://doi.org/10.1016/0006-8993(81)90078-0

Palmer, A. R., & Russell, I. J. (1986). Phase-locking in the cochlear nerve of the guinea-pig and its relation to the receptor potential of inner hair-cells. *Hearing Research*, *24*(1), 1–15. https://doi.org/10.1016/0378-5955(86)90002-X

Pfingst, L. X. E., Arbor, A., & Pfingst, B. E. (2003). Relative importance of temporal envelope and fine structure in lexical-tone perception Relative importance of temporal envelope and fine structure in lexical-tone perception (L), *3024*(L). https://doi.org/10.1121/1.1623786

Qin, M. K., & Oxenham, A. J. (2005). Effects of Envelope-Vocoder Processing on F0 Discrimination and Concurrent-Vowel Identification.

Rose, J. E., Brugge, J. F., Anderson, D. J., & Hind, J. E. (1967). Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey. *Journal of Neurophysiology*, *30*(4), 769–93. https://doi.org/10.1152/jn.1967.30.4.769

Rosen, S. (1992). Temporal information in speech: acoustic, auditory and linguistic aspects. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *336*(1278), 367–73. https://doi.org/10.1098/rstb.1992.0070

Schimmel, S., & Atlas, L. (2005). Coherent Envelope Detection for Modulation Filtering of Speech. In *Proceedings. (ICASSP '05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005.* (Vol. 1, pp. 221–224). IEEE. https://doi.org/10.1109/ICASSP.2005.1415090

Shannon, R. V., Zeng, F.-G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech Recognition with Primarily Temporal Cues. *Science*, *270*(5234).

Sheft, S., Ardoint, K., Lorenzi, C., & Sheft, S. (2008). Speech identification based on temporal fine structure cues Speech identification based on temporal fine structure cues, *562*. https://doi.org/10.1121/1.2918540

Siegenthaler, B. M. (1969). Maturation of Auditory Abilities in Children. *International Audiology*, *8*(1), 59–71. https://doi.org/10.3109/05384916909070190

Smith, Z. M., Delgutte, B., & Oxenham, A. J. (2002). Chimaeric sounds reveal dichotomies in auditory perception. *Nature*, *416*(6876), 87–90. https:// doi.org /10.1038/416087a

Soli, D. J. V. T. D., Kirby, V. M., Widin, G. P., Tasell, D. J. Van, Soli, S. D., Kirby, V. M., & Widin, G. P. (1998). Speech waveform envelope cues for consonant recognition, *1152*(1987). https://doi.org/10.1121/1.395251

Stone, M. A., Füllgrabe, C., & Moore, B. C. J. (2008). Benefit of high-rate envelope cues in vocoder processing: Effect of number of channels and spectral region. *The Journal of the Acoustical Society of America*, *124*(4), 2272–2282. https://doi.org/10.1121/1.2968678

Stone, M. A., & Moore, B. C. J. (2003). Effect of the speed of a single-channel dynamic range compressor on intelligibility in a competing speech task. *The Journal of the Acoustical Society of America*, *114*(2), 1023–1034. https://doi.org /10.1121 /1.1592160

Voelcker, H. B. (1966). Toward a unified theory of modulation part I: Phase-envelope relationships. *Proceedings of the IEEE*, *54*(3), 340–353. https://doi.org/10.1109 /PROC.1966.4695

Walker, G., Byrne, D., & Dillon, H. (1984). The effects of multichannel compression/ expansion amplification on the intelligibility of nonsense syllables in noise. *The Journal of the Acoustical Society of America*, *76*(3), 746–757. https://doi.org /10.1121/1.391261

Walsh, E. G. (1960). *Experiment In Hearing. Quarterly Journal of Experimental Physiology and Cognate Medical Sciences*. https://doi.org/10.1113/ expphysiol. 1960.sp001484

Xu, L., & Pfingst, B. E. (2003). Relative importance of temporal envelope and fine structure in lexical-tone perception (L). *The Journal of the Acoustical Society of America*, *114*(6), 3024–3027. https://doi.org/10.1121/1.1623786

Xu, L., Thompson, C. S., & Pfingst, B. E. (2005). Relative contributions of spectral and temporal cues for phoneme recognition. *The Journal of the Acoustical Society of America*, *117*(5), 3255–3267. https://doi.org/10.1121/1.1886405

Xu, L., & Zheng, Y. (2007). Spectral and temporal cues for phoneme recognition in noise. *The Journal of the Acoustical Society of America*, *122*(3), 1758–1764. https://doi.org/10.1121/1.2767000

Young, E. D., & Sachs, M. B. (1979). Representation of steady- state vowels in the temporal aspects of the discharge patterns of populations of auditory- nerve fibers. *The Journal of the Acoustical Society of America*, *66*(5), 1381–1403. https://doi.org/10.1121/1.383532

Zeng, F.-G., Nie, K., Liu, S., Stickney, G., Del Rio, E., Kong, Y.-Y., & Chen, H. (2004). On the dichotomy in auditory perception between temporal envelope and fine structure cues (L). *The Journal of the Acoustical Society of America*, *116*(3), 1351–1354. https://doi.org/10.1121/1.1777938