Perception of Some Temporal Parameters of Speech in Individuals with Auditory Dys-synchrony

DOCTORAL THESIS

Ajith Kumar U

Under the Guidance of

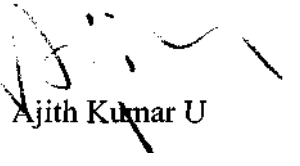
Prof. M Jayaram

Submitted to the University of Mysore
All India Institute of Speech and Hearing, Naimisham Campus, Manasagangothri, Mysore-5 70006

April, 2006

DECLARATION

I declare that this thesis entitled 'Perception of Some Temporal Parameters of Speech in Individuals with Auditory Dys-synchrony' which is submitted herewith for the award of the degree of Doctor of Philosophy in the Geld of Speech and Hearing to the University of Mysore, Mysore is the result of work carried out by me at the All India Institute of Speech and Hearing, Mysore, under the guidance of Prof. M. Jayaram, Director, All Institute of Speech and Hearing, Mysore. I further declare that the results of this work have not been previously submitted for any degree.



Ajith Kumar U

Place: Mysore

Date: 17.04.06

CERTIFICATE

This is to certify that the thesis entitled 'Perception of Some Temporal Parameters of Speech in Individuals with Auditory Dys-synchrony'. submitted by Ajith Kumar U, for the degree of Doctor of Philosophy in Speech and Hearing to the University of Mysore, was carried out at the All India Institute of Speech and Hearing, Mysore under my guidance.

Place: Mysore
Date: 17. 04. 06

Prof. M Jayaram
Director,
AIISH, Mysore

CERTIFICATE

This is to certify that the thesis entitled **'Perception of Some Temporal Parameters of Speech in Individuals with Auditory Dys-synchrony',** submitted by Ajith Kumar U for the degree of Doctor of Philosophy in Speech and Hearing to the University of Mysore was carried out at the All India Institute of Speech and Hearing, Mysore.

Place: Mysore
Dale: 17.04. 06

Prof. M. Jayaram
Director,
AIISH, Mysore

ACKNOWLEDGEMENTS

This research work was supported by the AIISH fellowship awarded by All India Institute of Speech and Hearing, Mysore, India. The work was carried out at the Department of Audiology, All India Institute of Speech and Hearing, Mysore, India.

My deepest respect and gratitude are due to my guide and mentor, Prof. M. Jayaram, Director, All India Institute of Speech and Hearing, Mysore, India. He has been and continues to be a source of inspiration at every stage of this work, sharing his expertise thoughts and experience.

I thank H O D and faculty, Department of Audiology for their support

I thank all the subjects who participated in the study for their time and cooperation.

Last but not the least I thank all my friends, students and colleagues who supported me in every step.

CONTENTS

Papers Published

1) Prevalence and Audiological characteristics of Individuals with Auditory Dys-Synchrony
Paper accepted for publication in International Journal of Audiology

2) Auditory Processing in Auditory Nueropathy, (2005). Published in Brain Behavioral Functions, 27, 1-8

3) Processing and Perception of temporal modulations Through a Digital Hearing Aid. Manuscript under review in Ear and Hearing

4) Presentations

CHAPTER 1

INTRODUCTION

Ear is one of the most important links in the speech chain and is essential for communication. All information from the peripheral receptor organ (cochlea) is carried to the brain for analysis through afferent auditory pathways. Higher neurocenters have control over the peripheral receptors through efferent pathways (Huffman & Henson, 1990). Deficits anywhere in these structures and pathways will lead to hearing impairment. The major effect of hearing impairment is loss of some or all of the important acoustic cues which in turn affect speech perception and communication.

Different assistive devices from hearing aids to cochlear implants are available to facilitate hearing, and thereby communication, in individuals with hearing impairment. However, all hearing impaired individuals may not benefit from amplification devices. There is evidence to show that individuals with auditory dys-synchrony may not benefit from hearing aids (Starr, Picton, Sininger, Hood & Berlin,. 1996; Stein et al., 1996). The condition of auditory neuropathy is more recently called auditory dys-synchrony (Berlin, Hood & Rose, 2001). Persons with auditory dys-synchrony often complain that they hear, but that they do not understand. Furthermore, their problem in understanding speech is aggravated under listening situations where noise and reverberation is present to a greater degree than usual. The speech understanding deficits of individuals with

auditory dys-synchrony are disproportionate to the degree of their hearing loss unlike those in the cochlear hearing loss group (Li, Wang, Chen & Liang, 2005; Starr et al., 1996).

1.0 Audiological Profile In Auditory Dys-synchrony

The disorder of auditory dys-synchrony is characterized by absence of auditory brainstem responses (ABR) in the presence of normal otoacoustic emissions and/or cochlear microphonics (Starr et al., 1996). However, if ABRs are present, then they are severely abnormal. In other words, a person with auditory dys-synchrony presents evidence of normal outer hair cell functioning, but abnormal auditory nerve functioning. In general, subjects who show evidence of (a) poor auditory functioning, and (b) poor auditory neural functions in the presence of normal outer hair cell functions are considered to have the condition of auditory dys-synchrony (Sininger&Oba,200l).

The general audiological findings in these patients suggest that responses which require intact auditory nerve or brainstem pathways like the acoustic reflex, ABRs, masking level difference and efferent suppression of otoacoustic emissions are abnormal. As said earlier, ABRs are generally absent, but when present, are severely abnormal. The extent of abnormality is disproportionate to the subject's audiometric thresholds for puretones. On the other hand, cochlear responses like otoacoustic emissions and cochlear

microphonics that are a result of normally functioning outer hair cells are normal (Berlin, 1999; Santarelli & Arslan, 2002; Starr et al., 1996).

1.1 Pathophysiology of Auditory Dys-synchrony

Auditory dys-synchrony may affect the functioning of inner hair cells, synaptic junctions between the inner hair cells and auditory nerve, or the auditory nerve itself (Starr et al., 1996). The physiologic changes accompanying auditory dys-synchrony may lead to a reduction in the number of conducting fibers due to axonal loss. In an histopathological study of cochlea and auditory nerve in an individual with auditory dys-synchrony, Starr et al. (2003) found that the organ of corti was normal throughout the cochlea except at the apical turn where a 30% loss of outer hair cells was observed. But, the inner hair cells were normal throughout the length of the cochlea. However, there was a profound loss of ganglion cells (>95%). The auditory nerve adjacent to the cochlear nucleus showed a marked reduction in the number of auditory fibers. Furthermore, the myelin sheath on the surviving auditory nerve fibers was thin indicating incomplete remyelinization. Therefore, it may be said that reduced neural input due to axonal loss could result in the loss of acoustic reflexes, that is, middle ear muscle and olivocochlear reflexes (Starr, Picton & Kim, 2001). These reflexes are evoked using suprathreshold signals that activate large number of auditory nerve fibers and this may not be possible in persons with auditory dys-synchrony because of large axonal loss.

Persons with auditory dys-synchrony may also manifest asynchronous firing of the auditory nerve fibers due to demyelinization (Starr et al., 2001). Demyelinization affects salutatory conduction and thereby slows down the conduction velocity of the nerve fibers. If the extent of slowing varies from one fiber to next (due to different degrees of demyelinization), then it leads to temporal asynchrony in the firing of the auditory nerve fibers thereby reducing the compound action potential of the auditory nerve. Asynchronization not only affects ABRs, but also influences auditory perception dependent on temporal cues (Kraus et al., 2000; Rance, Mckay, & Grayden, 2004; Starr et al., 1991; Starr et al., 1996: Zeng, Oba, Garde, Sininger, & Starr, 1999; Zeng, Kong, Michalewsk, & Starr, 2005). Axonal loss and demyelinization can occur together.

1.2 Speech Perception in Auditory Dys-synchrony

It is now well established that speech identification abilities of individuals with auditory dys-synchrony are disproportionate to the degree of their hearing loss (Li, et al., 2005; Starr, et al., 1996). In fact, it can be said that disproportionately poor speech identification scores which is disproportionate to puretone thresholds is the cardinal characteristic of persons with auditory dys-synchrony. Speech perception abilities in these patients appear to depend on the extent of distortion of temporal cues at suprathreshold levels rather than access to speech spectrum (related to audibility), unlike in patients with cochlear hearing loss (Ranee et al., 2004;

Zeng et al., 1999; Zeng et al., 2005). Zeng et al. (1999) reported abnormal results on two measures of temporal perception in a group of children with auditory dys-synchrony: (i) gap detection threshold (identification of silence embedded within bursts of noise) and (ii) temporal modulation transfer function (measure of sensitivity to slow and fast amplitude fluctuation). They also found a correlation between temporal modulation transfer function (TMTF) and speech identification scores in their patients. They modified temporal parameters of speech to a level which produced temporal processing deficits in persons with auditory dys-synchrony. They administered such edited speech to normal hearing listeners and found that the edited acoustic samples resulted in poor speech recognition abilities in normals similar to those seen in individuals with auditory dys-synchrony. Zeng et al. (1999) concluded that asynchronous firing of the auditory nerve resulted in distorted temporal coding of speech which in rum resulted in poor speech recognition that was disproportionate to the degree of hearing loss. Ranee et al. (2004) also reported poor performance on tasks involving timing cues (TMTF, temporal aspects of frequency discrimination) in a group of 14 children with auditory dys-synchrony. Specifically, processing abnormalities on these temporal tasks were significantly correlated with speech identification scores. Ranee et al. attributed these disproportionate speech identification scores to deficits in the processing of temporal information.

It has been shown that performance of individuals with auditory dys-synchrony is similar to normals on perception of intensity related

information such as sound localization based on interaural level difference and loudness discrimination (Zeng et al., 2005). In contrast, the subjects of Zeng et al. (2005) exhibited severe problems in timing related perception like temporal integration, gap detection, temporal modulation detection, backward and forward masking, and sound localization using interaural time differences. Individuals with auditory dys-synchrony have difficulty in detecting short duration acoustic signals, but not longer ones.

Another factor that is reported to be related to speech identification scores in individuals with auditory dys-synchrony is cortical event-related potentials. Ranee, Con-Wesson, Wunderlich and Dowell (2002) reported that a subgroup of children with auditory dys-synchrony, who had recordable cortical evoked potentials, performed well on an open set speech identification task and derived significant benefit from amplification. In contrast, subjects who had no recordable cortical evoked potentials performed poorly on speech identification tasks. Ranee et al. concluded, on the basis of this observation, that presence of cortical auditory evoked potential reflects some amount of preserved synchrony in central auditory system which contributes to better speech understanding despite the distortion that occurs at 8^ nerve and auditory brainstem levels in these individuals.

Only limited data is available on fine-grained speech perception abilities of individuals with auditory dys-synchrony. Kraus et al. (2000) presented data on electrophysiological, psychophysical and speech perception characteristics in a rare case of auditory dys-synchrony who had normal hearing thresholds. They studied just noticeable differences for

synthesized consonant-vowel continua /ba-wa/ and /da-ga/. The subject showed poor performance on /da-ga/ continuum. In addition, presence of noise degraded the discrimination ability of this subject to a greater extent than it did in normals. Kraus et al. concluded that timing information, at stimulus onset, was most vulnerable to disruption while representation of long duration steady state timing cues was better preserved.

1.3 Need for Studies on Auditory Dys-synchrony

It was generally believed, until 1995, that auditory dys-synchrony is a rare disorder. On the contrary, it is not such an extremely rare disorder. The reported prevalence of auditory dys-synchrony varies from 11% to 0.5% (Davis & Hirsh, 1979; Ranee, et al., 1999; Tang, McPherson, Yuen, Wong, & Lee, 2004). Davis and Hirsh (1979) reported that 1 in 200 hearing impaired children exhibit an audiological picture that is consistent with the contemporary diagnosis of auditory dys-synchrony. Ranee et al. (1999 - hospital based statistics) assessed 5199 'at risk' children for auditory dys-synchrony. The prevalence of children with auditory dys-synchrony in at risk population was 1 in 433 (0.23%) and in children with hearing impairment was 1 in 9 (11.01%). Tang et al. (2004) investigated the frequency of occurrence of auditory dys-synchrony in school-aged hearing-impaired children and reported a prevalence of 2.44%. However, there is no information on the prevalence of this problem in the Indian scenario. If the prevalence of this problem in our population is comparable with that

reported in the Western population, then further studies on this population is justified.

1.4 Statement of the Problem

Previous studies on speech perception and psychophysical abilities of individuals with auditory dys-synchrony suggest markedly affected temporal processing in these individuals. In fact, there is reasonable evidence to show that persons with auditory dys-synchrony find it more difficult to perceive short duration dynamic sounds than long duration steady sounds (Kraus et al., 2000; Zeng, et al., 2005). Therefore, the purpose of this study is to understand the role of some short duration temporal parameters of speech segments on the perception of speech in individuals with auditory dys-synchrony. Specifically, this study aimed at investigating perception of speech in individuals with auditory dys-synchrony when temporal parameters of speech like transition duration, burst duration and voice onset time were modified. Related to this, the study also investigated the influence of these temporal parameters, in isolation and in combination, modified in the units of JNDs, on speech perception. Furthermore, temporal resolution in individuals with auditory dys-synchrony was studied by measuring TMTF and its relation to speech perception. Another issue that this study addressed related to the prevalence of the problem of auditory dys-synchrony in Indian population. This was done to see if the extent of

the problem is large enough to justify focused research into auditory dys-synchrony.

1.5 Objectives of the Study

The objectives of this study were to

a) find the prevalence of the problem of auditory dys-synchrony in Indian population and to describe general audiological characteristics in this clinical population,

b) measure the just noticeable difference for voice onset time, transition duration, and burst duration of speech segments in individuals with auditory dys-synchrony and to compare these with those for normal listeners,

c) investigate the effect of changed voice onset time, transition duration, and burst duration of speech segments, in isolation and in combination, on speech perception in individuals with auditory dys-synchrony, and

d) obtain TMTF by measuring detection thresholds for sinusoidally amplitude modulated white noise at 4 Hz, 16 Hz, 32 Hz, 64 Hz, 128 Hz and 200 Hz and compare these results with speech perception abilities in persons with auditory dys-synchrony.

1.6 Justification for the Study

Patients with auditory dys-synchrony have marked deficits in processing temporal information, but spared ability for processing intensity and frequency information (Ranee et al., 2004). This is in contrast to patients with cochlear hearing loss who typically demonstrate loudness recruitment and broadening of auditory filters, but, normal processing of temporal information, at least at high sensation levels. Persons with cochlear hearing loss derive significant benefit from hearing aids which employ nonlinear compression circuits. All these hearing aids assume abnormal functioning of outer hair cells (Berlin, Hood, Hurely & Wen 1996). Hence, these aids are of not much use for individuals with auditory dys-synchrony who have normal outer hair cell functioning. Other management strategies used by individuals with auditory dys-synchrony include FM systems, cochlear implants, perceptual training, speech reading and cued speech (Kraus, 2001). Cochlear implant may be a viable option for some of auditory dys-synchrony patients. Many studies have shown that benefit from cochlear implants for individuals with auditory dys-synchrony is comparable to that of individuals with cochlear hearing loss (Buss, Labadie, Brown, Gross, Grose & Pillsbury, 2002; Miyamoto, Kirk, Renshaw, & Hussain, 1999; Peterson, et al., 2003; Shallop, Peterson, Facer, Fabry, & Driscoll, 2001; Trautwein, Sininger & Nelson, 2000). But, the degree of hearing loss may not justify the use of cochlear implant in some of these individuals with auditory dys-synchrony as an option of management.

Therefore, a cochlear implant for persons with auditory dys-synchrony is debatable. Hence, there is a need to explore alternative strategies of speech processing that are much less expensive and invasive than cochlear implants, but, which benefit individuals with auditory dys-synchrony.

The available hearing aids do not faithfully reproduce the stimulus particularly the rapid temporal changes. This is especially at true at high modulation frequencies. The result is that certain amount of distortion is introduced in to the processed signal and this distortion may interfere with speech perception (Kumar & Jayaram, 2005). Hence, a new type of hearing aid design is called for to improve speech recognition in patients with auditory dys-synchrony. The new improved design should provide for augmenting/modifying temporal information of speech. It would be necessary to understand the effect of modification of temporal aspects of speech on their perception in patients with auditory dys-synchrony to develop and implement an alternate speech processing algorithm in hearing aids. Patients with auditory dys-synchrony show speech recognition abilities, both in quite and noise, that are disproportionate to their degree of hearing loss for puretones.

Several psychophysical studies, have shown using nonspeech stimuli, marked deficits in processing temporal information in individuals with auditory dys-synchrony while frequency and intensity coding are only marginally affected. It can be hypothesized that these temporal processing deficits distort speech leading to poor speech recognition in these individuals. Hence, it is necessary to study the temporal processing deficits

and effects of modification of temporal parameters of speech, on speech perception in individuals with auditory dys-synchrony.

CHAPTER 2

REVIEW

The term auditory dys-synchrony has been used to describe a form of hearing impairment in which cochlear amplification function is normal, but neural transmission in the auditory pathways is disordered. This clinical entity was first described in detail by Starr et al. (1991) in one subject and was thought to include a dysfunction of the auditory nerve. Subsequently, ten subjects with similar symptoms were identified (Starr, Picton, Sininger, Hood & Berlin, 1996). As eight of these subjects had accompanying peripheral neuropathies, this condition was christened auditory neuropathy. However, more recently, Berlin, Hood and Ross (2001) suggested the term auditory dys-synchrony as a more accurate indicator of the underlying condition. The coinage of the term was based on two primary underlying factors:

a) the auditory nerve itself may not be affected, and

b) the term auditory neuropathy may lead clinicians not to consider cochlear implant as a management option. It has subsequently been shown that these individuals with auditory dys-synchrony benefit from cochlear implants (Berlin, Hood, Morlet, Rose, & Brashears, 2003; Berlin, Morlet, & Hood, 2003; Peterson et al., 2003).

It is generally agreed upon by audiologists that a person to be diagnosed with auditory dys-synchrony must have the following features (Sininger & Oba, 2001):

a) evidence of poor auditory functioning - patient must be having difficulty in understanding speech at least in some situations regardless of the level of pure tone hearing thresholds, and

b) evidence of poor neural functioning - patients should not have ABRs, and if present, they should be abnormal as well as elevated. The acoustic auditory brainstem reflexes (stapedial reflex, medial olivocochlear bundle reflex and masking level difference) are generally not present, and

c) evidence of normal cochlear amplification function - patients must show either cochlear microphonics or otoacoustic emissions.

1. 0 Clinical Profile of Patients with Auditory Dys-Synchrony

1.1 Onset and Course

There is not much research information available on the age of onset of auditory dys-synchrony. However, evidence available points to childhood as the age at which onset takes place. Sininger and Oba (2001) studied a group of 59 individuals with auditory dys-synchrony and reported a mean age of onset of 9 years. 75% of their patients were less than 10 years of age when the first symptom of auditory dys-synchrony was seen. In fact, the age of onset of auditory dys-synchrony in the clinical population of Sininger and Oba ranged between birth to 60 years of age with the largest group reporting onset of the problem before 2 years of age. As the problem of auditory dys-synchrony

recovers after the onset, or fluctuates (Berlin, 1999), the true age of the onset of the problem may never be known.

Equally unpredictable is the course of auditory dys-synchrony. The condition resolves or remains the same in some individuals. Hearing thresholds fluctuate in some, and get worse in some. Berlin (1999) identified several patterns in the time course of auditory dys-synchrony. Some of these are as follows:

a) Some patients show retrograde loss of cochlear microphonics and otoacoustic emissions and become almost indistinguishable from patients with cochlear hearing loss.

b) Some patients retain cochlear microphonics and otoacoustic emissions, but cannot leam speech - language by auditory mode alone. Visual information (cued speech, sign language, speech reading) is necessary.

c) Some patients show worsening of the symptoms and develop other neuropathies. Starr, Sininger and Praat (2000) reported peripheral neuropathy in 40% of their patients with auditory dys-synchrony. None of the children below five years of age in their group showed clinical evidence of peripheral neuropathy whereas 80% of the patients examined after the age of 15 years showed both clinical and physiological (nerve conduction) evidence of peripheral neuropathy. Hence, Starr et al. hypothesized that some of the patients may subsequently develop peripheral neuropathies as they grow.

d) Some patients loose their otoacoustic emissions, but not cochlear microphonics. Such patients may manifest severe hearing impairment,

while occasionally, they may also show evidence of unexpected hearing abilities (Withnell, 2001). Around 20% of the individuals with auditory dys-synchrony loose their otoacoustic emissions as the disease progresses (Starr, et al., 2000).

e) Some patients go through life without complaining of any problem. They develop speech and language normally and would be identified as cases of auditory dys-synchrony only if an ABR had to be done for some reason (either as a part of a screening procedure or a research project).

f) Some patients with auditory dys-synchrony may show fluctuations in their hearing abilities that are temperature sensitive. Children have been reported to show variation in puretone thresholds, presence or absence of ABRs etc. depending on whether they were febrile or afebrile (Starr, sininger, Winter, Derebery, Oba, & Michalewski, 1998). In fact, many patients with auditory dys-synchrony appear to experience moment-to-moment fluctuations in their hearing sensitivity which may be misinterpreted as inconsistent response during testing.

1.2 Prevalence

The extent of prevalence of auditory dys-synchrony is still not clearly known. The reported prevalence rate varies from 11%, to 0.5% (Davis & Hirsh, 1979; Kraus, Ozdamar, Stein & Reed 1984; Ranee, et al., 1999; Tang, Mcpherson, Yuen, Wong, & Lee, 2004). Many studies have reported

electrophysiological fmdings, in groups of patients, which are consistent with the contemporary diagnosis of auditory dys-synchrony (Davis &Hirsh, 1979; Kraus, et al., 1984). All such subjects may or may not turn out to be cases of auditory dys-synchrony, and hence the uncertainty when one talks of prevalence of auditory dys-synchrony. Prevalence studies of auditory dys-synchrony have been carried out on 'at risk' children (Ranee et al., 1999) and school-aged hearing impaired children (Tang et al., 2004). Tang et al. reported a prevalence of 2.44% in school-aged hearing impaired children. Ranee et al. in a study of 5199 at risk children, reported a prevalence of I in 433 (0.23%) while prevalence in children with hearing impairment was I in 9 (11.01%). A large number of subjects in the study of Ranee et al. had been bom premature and were at risk for neurodevelopmental disorder. The risk factor identified, in half of this population with auditory dys-synchrony, was hyperbilirubinemia. It is known that hyperbilirubinemia can cause both permanent and temporary dysfunction of auditory pathways (Gupta, Raj, Anand, 1990). Some of the infants may recover and may show ABR as they grow.

1.3 Audiological Characteristics

1.3.1 Puretone and Speech Identification Profile

Audiological profile of persons with auditory dys-synchrony is variable. Auditory dys-synchronics present all levels of hearing loss. The

degree of hearing loss may range anywhere from normal healing sensitivity to profound loss. Majority of the patients show bilateral symmetrical hearing loss (Ranee et al., 1999; Swinger & Oba, 2001). However, it is difficult to determine the degree of hearing loss in persons with auditory dys-synchrony as they not only show inconsistent responses, but many also show reverse sloping or peaked audiograms.

Rance et al. (1999) noted that subject's audiometric configuration varied with the degree of hearing loss. Ears with normal or near normal hearing acuity showed equal sensitivities at all the frequencies. Subjects with mild to severe hearing loss had audiograms that showed poor hearing sensitivities in the low and mid frequencies, but better thresholds in the high frequencies. Starr et al. (2000), in a study of 67 patients with auditory dys-synchrony, reported flat audiogram in 41%, reverse sloping audiogram in 29%, an irregular saw-tooth pattern in 9%, a 'U' shaped audiogram in 5%, and a tent shaped audiogram with a peak usually at 2 kHz in 5% of the patients. Only 11% had high frequency sloping which is typical of cochlear hearing loss. However, 43% of the patients of Sininger & Oba (2001) showed flat audiometric shape while 28% had reverse sloping loss with higher thresholds at low frequencies than at high frequencies.

Reverse sloping audiogram seen in patients with auditory dys-synchrony is further evidence that the underlying etiology of hearing loss in auditory dys-synchrony is neural and not cochlear. These reverse sloping audiograms also indicate that individuals with auditory dys-synchrony have problems in processing phase locked temporal information. According to

theories of pitch perception (see Moore 2003, for a detailed discussion), high frequencies above 1 kHz are signalled by place of excitation on the basilar membrane. This is because auditory nerve fibers cannot fire at higher rates due to limitations imposed by refractory periods. The auditory nerve fibers can fire for each phase of the signal for low frequency puretones, but not for high frequency tones. So, poor low frequency thresholds indicate poor auditory percepts dependent on temporal cues (Starr, Picton & Kim 2001).

It has been noted that patients with auditory dys-synchrony have speech perception abilities that are out of proportion with their pure tone hearing loss (Li, Wang, Chen, & Liang, 2005; Starr, et al., 1996). Also, speech perception abilities of persons with auditory dys-synchrony is highly variable, with some patients performing at levels expected for cochlear hearing loss of the same degree, while some others show little or no measurable speech identification despite having adequate sound detection abilities. Furthermore, this discrepancy between sound detection and speech identification appears to be related to suprathreshold distortion of temporal cues rather than audibility (Ranee, Mckay, & Grayden, 2004; Zeng, Oba, Garde, Sininger, & Starr, 1999; Zeng, Kong, Michalewsk, & Starr, 2005).

1.3.2 Physiological and Electrophysiological Responses

Auditory brainstem responses are generally absent in persons with auditory dys-synchrony. However, if present, then they are severely abnormal. Starr et al. (2000) reported that 70% of their patients did not show

any component of ABR regardless of the level of the stimulus. 19% of the remaining showed abnormal wave V, and in most of them, the peak was clearly defined though characterized by abnormal amplitude and latency. 6% did show wave III and V, but their amplitude, wave morphology and latency was abnormal. A common feature in patients who showed ABR was the increased sensitivity of the response to an increase in stimulation rate. At high stimulation rates, none of the patients with auditory dys-synchrony showed any ABR component. This increased sensitivity to rate may be due to decreased synaptic efficiency which becomes apparent at high stimulation rates. Generally, in patients with auditory dys-synchrony, behavior thresholds are not correlated with ABR thresholds (Starr et al., 1996; Starr et al., 2000).

Many patients with auditory dys-synchrony demonstrate long duration cochlear microphonics which may be confused with early peaks of ABR. Therefore, it is recommended that in patients who show ABR, responses for the condensation and rarefaction clicks should be compared to rule out the possibility of misinterpreting cochlear microphonics as ABR (Berlin et al., 1998). Otoacoustic emissions are present in most individuals with auditory dys-synchrony. Starr, et al. (2000) reported that about 80% of the patients with auditory dys-synchrony had clear OAEs. 9% of the subjects did not show OAEs in the initial evaluation while OAEs disappeared over time in 11 to 16% of patients. Cochlear microphonics represent outer hair cell activity in the basal region. Cochlear microphonics can be recorded from individuals with auditory dys-synchrony. They are robust and present for several milliseconds after the transient click (Berlin, 1999; De1tenr, et_al., 1998; Duan & Wang,

2002; Starr, et al., 1996; Starr, et al., 2000; Santarelli & Arslan, 2002). Auditory brainstem acoustic reflexes (contralateral suppression of OAE, acoustical stapedial reflex and masking level difference) are absent in persons with auditory dys-synchrony. The typical audiological profile of persons with auditory dys-synchrony is summarized in Table 2.1.

Table 2.1. Audiological profile of persons with auditory dys-synchrony.

Audiological tests	Result
Degree of hearing loss	Normal to profound
Configuration of audiogram	Variable, mostly reverse sloping
Tympanogram	Normal
Acoustic reflexes	Absent
Non acoustic reflexes	Present
ABR	Absent or severely abnormal, if present
OAE	Usually present
Contralateral suppression of OAE	Absent
Cochlear microphonics	Present
Masking level difference	Absent

2.0 Pathology

Starr et al. (2001) examined the sural nerve in 6 patients with auditory dys-synchrony and a concomitant peripheral neuropathy. Figure 2.1 shows photon micrographs of a cross section of the sural nerve of one of the subjects. Biopsy showed axonal neuropathy which had been responsible for the loss of large myelinated fibers. Three subjects had axonal loss with evidence of secondary demyelinization as well as remyelinization of remaining fibers. Biopsy of the sural nerve of two other subjects showed extensive loss of both axons and myelin sheath. Starr et al. (2003) reported, in an histopathological study of the cochlea and auditory nerve of a patient with auditory dys-synchrony, that the organ of corti was normal throughout the cochlea except in the apical turn where a 30% loss of outer hair cells was found. The inner hair cells were normal throughout the length of the cochlea. However, there was a profound loss of ganglion cells (>95%). There were just 1161 and 1548 surviving ganglion cells in the right and the left ear, respectively. The age appropriate normal count for ganglion cells is 23,000. The auditory nerve adjacent to the cochlear nucleus showed a marked reduction in the number of auditory fibers. Furthermore, the myelin sheath on the surviving auditory nerve fibers was thin indicating incomplete remyelinization. Consequences of these two pathologies, that is, loss of ganglion cells and demyelinization to temporal processing of acoustical signal is discussed in the following section.

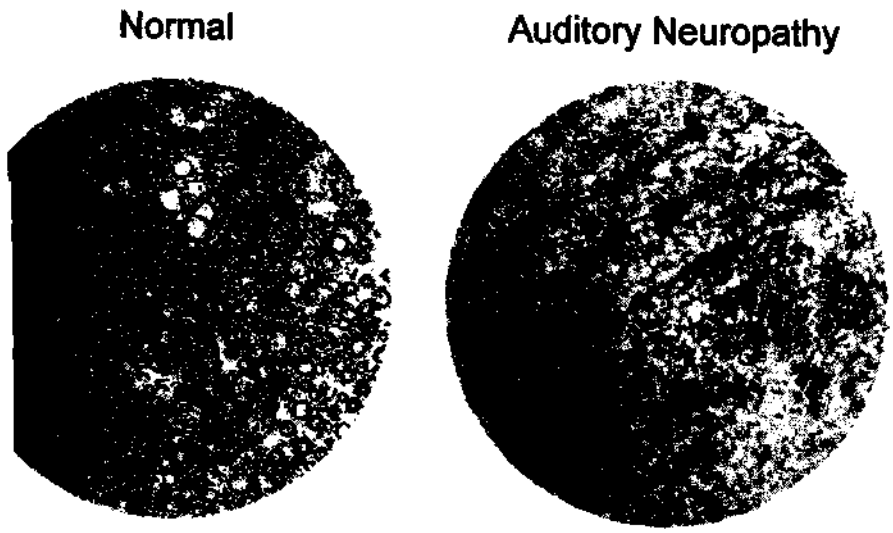


Figure 2.1. Photon micrographs of a cross section of the sural nerve of a patient with auditory dys-synchrony and normal hearing subject. Note that there is marked reduction and demyelination of nerve fibers. (From Auditory neuropathy, A new perspective on hearing disorders by Sininger and Starr (2001). Reprinted with permission from Delmar Learning, a division of Thomson Learning).

2.1 Pathophysiology of Auditory Dys-synchrony

Neuropathies may be caused by a primary demyelination or by an axonal disease. Demyeilinating neuropathies affect the Schwan cells which form the myelin sheath around the axons. Myelin plays an important role in the rapid salutatory transmission of nerve impulses along the axons. Demyelinating neuropathies slowdown or block nerve conduction and produce motor or sensory symptoms distal to the site of demyeilinization. Demyelinated fibers are poor conductors of rapid trains of action potentials. If demyelination affects all the auditory nerve fibers to the same degree, then

transmission through all fibers will be slowed down, and amplitude of the compound action potential will be unaffected despite slowing of conduction velocity (Figure 2.2, second column). On the other hand, if the extent of slowing varies from one fiber to the next, then the amplitude of action potential becomes small and smeared (Figure 2.2, third column). This smeared temporal representation of the acoustic stimulus may influence auditory perception that is dependent on temporal cues.

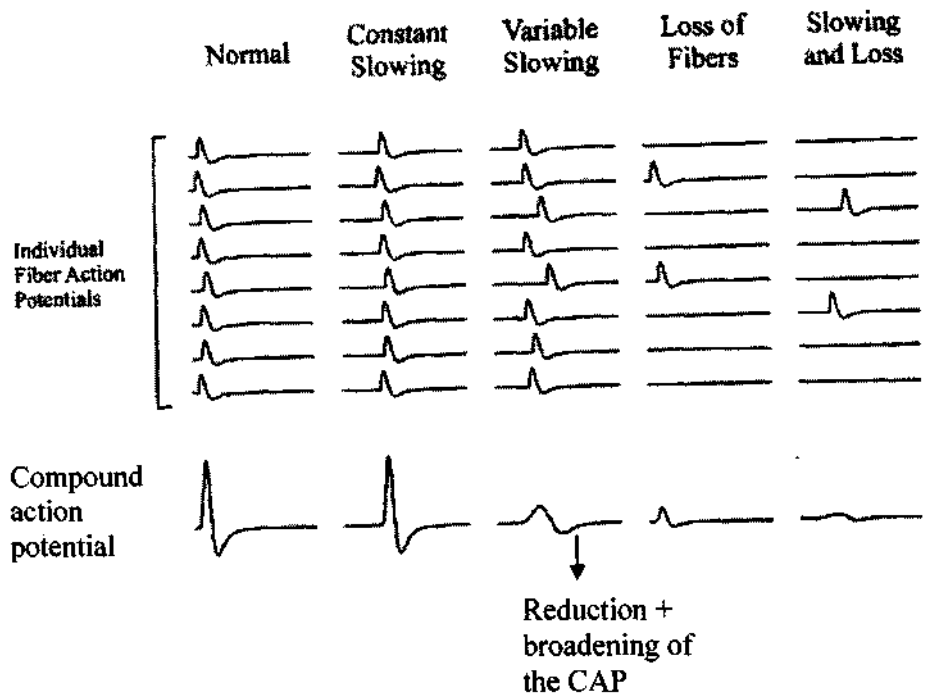


Figure 2.2. Action potential of individual fibers and resultant compound action potential. (From Auditory neuropathy, A new perspective on hearing disorders by Sininger and Starr (2001). Reprinted with permission from Delmar Learning, a division of Thomson Learning).

As the name suggests, axons are affected in axonal neuropathy. The hallmark of many axonal neuropathies is a retrograde degeneration of the distal portion of peripheral nerves. This is because of inadequate transport of metabolic substrate and growth factors between neuronal cell body and distal portion of its axon. In this condition, though the nerve fibers are reduced in number they function normally in terms of speed of conduction. As the number of the nerve fibers is reduced, a resultant compound action potential is reduced in amplitude.

Higher discharge rate in response to intense acoustic stimuli contributes to reflex activation of middle ear muscles. Acoustic reflexes of middle ear muscles are typically absent in subjects with auditory dys-synchrony and the reason for this may be the failure of the auditory nerve fibers to develop sufficiently higher discharge rates to activate the motor neurons of stapedial muscle. The longer fibers are more susceptible to axonal neuropathies. The longest cochlear nerve fibers are those innervating the apex of the cochlea (which mediates low frequencies). The shortest fibers are those innervating the second half of the first cochlear turn (which mediates middle frequencies). Those innervating the basal parts of the cochlea (mediating high frequencies) have lengths in between these two extremes (Starr et al., 2001). If persons with auditory dys-synchrony were to exhibit a 'dying back' pathophysiology, then the mid frequencies must be affected to a lesser extent than the low and high frequencies. This was looked in to in the present study in retrospective analyses of the audiometric configuration of 61 individuals with auditory dys-synchrony.

3.0 Psychophysics and Speech Perception in Patients with Auditory Dys-synchrony

3.1 Intensity and Frequency Processing

Zeng et al. (2001) analyzed the loudness growth function in one subject with auditory dys-synchrony using magnitude estimation and loudness

scaling techniques. Results showed that the subject demonstrated a larger compressive loudness function than the normal control subject. In another study, Rance et al. (2004) demonstrated that persons with auditory dys-synchrony show a slightly larger difference limenat low sensation levels than normals, but it approached normal values at high sensation levels similar to that seen in normals.

Rance et al. (2004) further demonstrated that frequency resolution in individuals with auditory dys-synchrony, measured using notched-noise technique, was normal. This finding is consistent with the presence of normal functioning outer hair cells in individuals with auditory dys-synchrony. It is well established that normal functioning outer hair cells are required for good frequency resolution in the cochlea (Dallos & Corey, 1991). Frequency discrimination ability of patients with auditory dys-synchrony is significantly poorer compared to that of normal hearing subjects, particularly at low frequencies (Rance et al., 2004; Starr et al., 1991; Starr et al., 1996; Zeng et al., 2005). Rance, et al. measured the frequency difference limen and frequency modulation detection limen for 500 Hz and 4 kHz in children with auditory dys-synchrony. Results showed that both the difference limens were better at 4 kHz compared to those at 500 Hz. Furthermore, frequency difference limens were better compared to frequency modulated difference limen scores. Also, frequency discrimination abilities were strongly correlated with speech perception scores. Similar results were reported by Zeng et al. (2005). Zeng et al. reported that individuals with auditory dys-synchrony demonstrated greater impairment in frequency discrimination at low

frequencies (< 4000 Hz) than at high frequencies. These results can be explained on the basis of differential mechanisms of frequency coding at high and low frequencies. Frequency discrimination is thought to be dependent on spatial changes in the excitation pattern along the basilar membrane for frequencies above 4 kHz (Sek & Moore, 1995). In contrast, discrimination of frequencies below 4 kHz is by the use of neural phase locking cues (Blackburn & Sachs, 1989; Goldberg & Brownell, 1973; Winter & Palmer, 1990). Individuals with auditory dys-synchrony may not use the phase locking cues to the same extent as normally hearing subjects do. Hence, high frequency discrimination that does not involve phase locking cues is relatively better compared to discrimination of low frequencies that depends on phase locking cues.

In general, these studies demonstrate that individuals with auditory dys-synchrony have relatively intact intensity and frequency resolution, but demonstrate impaired temporal aspects of frequency discrimination. Hence, it was felt that signal processing strategies that involve spectral enhancement/amplification, or filtering may not enhance speech perception abilities of persons with auditory dys-synchrony. Therefore, the focus of this study was to understand the effect of modification of three temporal parameters of TD, BD and VOT on speech perception in individuals with auditory dys-synchrony.

3.3 Temporal Processing

Several investigators have explored the temporal processing abilities of individuals with auditory dys-synchrony. Zeng et al. (2005) evaluated several time-related functions like temporal integration, gap detection, temporal modulation detection, backward masking, forward masking and simultaneous masking in individuals with auditory dys-synchrony. They adopted an adaptive three-interval three-alternative forced choice, two down one up procedure to track 70.7% correct response criterion. They found improvement in thresholds with increase in signal duration in individuals with auditory dys-synchrony as is the case with normals. However, the slope of the integration function was slightly elevated in individuals with auditory dys-synchrony (-9 dB per doubling of duration) than in normal hearing subjects (-3 dB per doubling of the duration). Similar results were reported by Starr et al. (1991). In contrast, Zeng et al. (1999, 2001) reported normal or near normal temporal integration functions in individuals with auditory dys-synchrony. This disagreement between the two studies cannot be because of methodological differences as the technique (psychophysical) and the stimulus used were the same in both the studies. The difference in statistical technique used might better account for the difference in result. While Zeng et al. (2005) compared results of individuals with auditory dys-synchrony with the mean and + ISE of normal control subjects, Zeng et al. (1999, 2001) compared the same with mean + 2SD of normal control data.

Abnormal gap detection (identification of silent period embedded within a noise burst) has been reported in individuals with auditory dys-synchrony (Zeng et al., 2005; Zeng et al., 1999, 2001). Zeng et al. (2005) reported poor gap detection thresholds in individuals with auditory dys-synchrony. Normal hearing individuals required a silent interval of around 50 ms to detect a gap at 5 dB SL. However, the detection threshold improved to 3 ms at higher sensation levels (30 to 40 dB SL). Individuals with auditory dys-synchrony performed similar to normal hearing subjects at low sensation levels, but unlike normals, required significantly larger gap to detect at higher sensation levels. This defies any explanation. However, Zeng et al. (1999, 2005) gave a phenomenological model to explain the abnormal gap detection thresholds in individuals with auditory dys-synchrony. This model assumes that the main effect of dys-synchronous activity is a smeared temporal representation of the acoustic stimulus (Figure 2.3). The sharp temporal changes in the physical representation of the stimulus are lost in the internal neural representation due to smearing of the waveform. If the listening task is to merely detect the sound, then this smeared representation will not affect perception. However, if the task is to discriminate between sounds - one with a gap and one without - then, smearing of the internal representation makes the task more difficult.

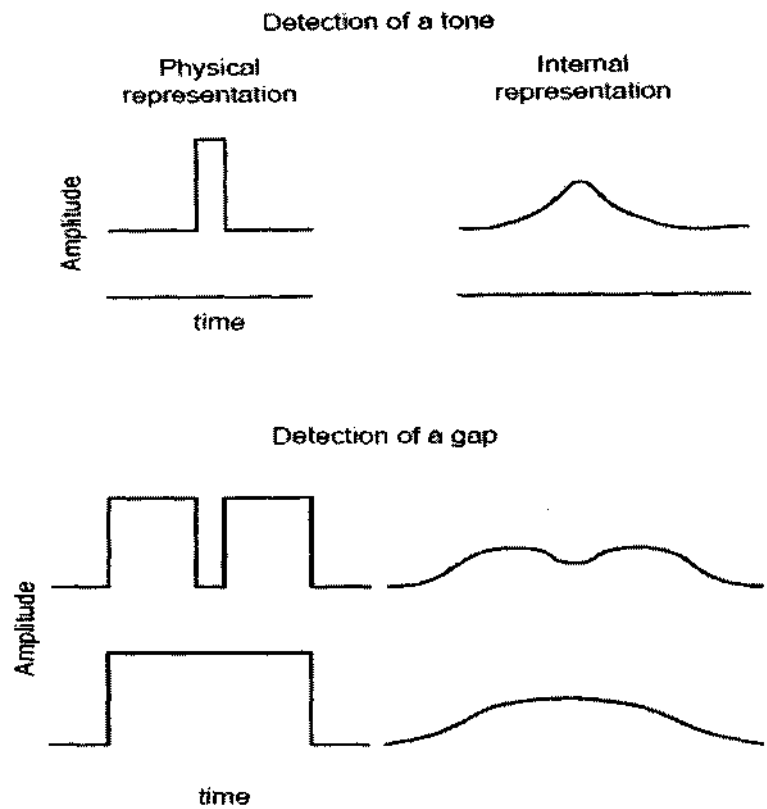


Figure 2.3. A phenomenological model of auditory dys-synchrony (Zeng et al., 1999). Smearing of the temporal envelope does not affect the detection of a tone (top panel) because this task requires an all-or-none decision. However, smearing causes major problem in gap detection (bottom panel) as the task requires finer discrimination of two different waveforms. (From Auditory neuropathy, A new perspective on hearing disorders by Sininger and Starr (2001). Reprinted with permission from Delmar Learning, a division of Thomson Learning).

Difficulty in detecting short duration acoustic events (as shown by psychoacoustic experiments on gap detection and temporal integration) may pose problems to persons with auditory dys-synchrony in processing brief, but critical events of speech. It is known that many critical elements for speech

perception such as transition, burst and VOT last only up to a few milliseconds. The studies of Zeng et al. (1999, 2001, 2005) have amply demonstrated that individuals with auditory dys-synchrony have severe problems in processing silent intervals of even 20-30 ms (comparable to lag VOT in unvoiced stops in Kannada language) as compared to normals and individuals with cochlear hearing loss. Also, it has been shown that thresholds for short duration signals are significantly higher (comparable to low energy lead VOT in voiced stops in Kannada language) in individuals with auditory dys-synchrony than in normals or individuals with cochlear hearing loss. These deficits may lead to difficulty in processing VOT. VOT is an important cue for the perception of voicing in Kannada and therefore, auditory dys-synchrony may have difficulty in processing voicing or differentiating voiced from unvoiced.

3.3.1 Temporal Modulation detection Thresholds

Another temporal process that has been reported to be abnormal in individuals with auditory dys-synchrony is the temporal modulation transfer function (Rance et al., 2004; Zeng et al., 1999; Zeng et al., 2005). Temporal modulation transfer function is a measure of sensitivity to amplitude fluctuation over a range of modulation frequencies. This measures one's ability to perceive changes in stimuli over time. Zeng et al. (1999) reported

that individuals with auditory dys-synchrony showed high peak sensitivity (-8.7 dB compared to -19.9 dB in normal controls) and lower cut off frequency (17 Hz compared to 258.1 Hz in normal controls). Rance et al. (2004) also reported significant differences in modulation detection thresholds between auditory dys-synchronics who have good and poor speech perception scores. Individuals with auditory dys-synchrony with speech identification scores less than 30% had poorer modulation detection thresholds compared to subjects who had more than 30% speech identification score.

Kraus et al. (2000) reported exaggerated masking effect in one patient with auditory dys-synchrony who had near normal hearing thresholds. Temporal masking and simultaneous masking paradigms have shown that individuals with auditory dys-synchrony have difficulty in separating sounds that occur successively as well as in detecting signal in noise (Zeng et al., 2005). In forward masking, individuals with auditory dys-synchrony showed 60% masking even when signal and masker were separated by as much as 100 ms while normal controls showed only 15% masking at a signal delay of <20ms. This shows that individuals with auditory dys-synchrony have difficulty in separating sounds that occur in close succession. This may produce difficulty in perceiving voice onset time, burst or transition in speech which are short acoustic events that occur in close succession and are critical for understanding speech. In simultaneous masking condition, individuals with auditory dys-synchrony showed excessive masking of about 20 dB compared to the normal control group (the difference was statistically significant). This excessive masking was independent of threshold at test

frequency which may be elevated or normal. The slope of the masking function varied between subjects - some having relatively normal slope, some having abnormally steep slope and still some others having abnormally shallow slope. Effect of masking was exaggerated when short duration signals were used. There was an 'over shoot' effect observed in simultaneous masking condition in normal hearing subjects. Over shoot' is defined as the difference in the detection threshold when a brief tone is placed at the onset of masker noise and later, in the steady portion of the noise. Normal subjects showed poorer thresholds when brief tone was placed at the onset of the noise. Similar results were found in auditory dys-synchrony group, but, with increased masking effect. Increased masking effects may have adverse effect on the perception of some important speech events like voice onset time, burst and transitions as these events may get masked by the succeeding or preceding vowel.

Experiments on modulation detection and masking have shown that individuals with auditory dys-synchrony have difficulty in perceiving the temporal envelop of the signal as well as in separating two closely occurring acoustic events. The importance of temporal envelop as well as the effects of smearing of temporal envelop on speech perception has already been demonstrated (Plomp, 1988; Kumar and Jayaram, 2005). Increased degree of temporal masking (backward and forward) may result in the masking of critical speech events (transition, burst and VOT) by the steady-state portion of the preceding or succeeding vowel. Effect of masking was more when short duration signals were used than when long duration signals were

employed. Hence, increasing the duration of some important short speech events might lead to better speech perception in individuals with auditory dys-synchrony by reducing the masking effects.

Zeng et al. (2005) also evaluated binaural processing of interaural time, interaural intensity, fusion and beats in individuals with auditory dys-synchrony. Subjects with auditory dys-synchrony, like normals, could effectively use interaural level difference (intensity) to localize sounds, but the performance of individuals with auditory dys-synchrony was significantly poorer compared to normals on experiments with intra aural time. Normal subjects localized the sound to the ear with leading phase while auditory dys-synchronics could not use interaural time cue to localize sounds. Subjects with auditory dys-synchrony performed similar to normal control group on monoaural beat task, but failed to perceive beats on binaural presentation. This could be because perception of monoaural beats requires spike synchrony to 3 Hz modulations in the waveform envelope, while detection of binaural beats require spike synchrony to rapidly changing carrier frequencies. This result shows that individuals with auditory dys-synchrony can perceive slow temporal fluctuations, but not the faster ones. Similar findings were reported by Starr et al. (1991) in a subject who fits the contemporary diagnosis of auditory dys-synchrony. Temporal processing abnormalities in auditory dys-synchronics have also been demonstrated using evoked potentials (Michalewski, Starr, Nguyen, Kong & Zeng, 2005). Evoked potentials are sensitive to temporal gaps. In normal subjects, evoked potentials (N100 and P200 components) can be recorded in response to gaps as short as 5ms. "Gap

evoked potentials" could be recorded in auditory dys-synchronics only for longer gap durations. There was a close association between gap detection thresholds measured psychoacoustically and electrophysiologically in both normal hearing subjects and individuals with auditory dys-synchrony (Zeng et al., 2005).

These psychophysical/electrophysiological findings point to the fact that timing and synchronicity in the firing of neurons in the auditory nerve as well as auditory brainstem regions are important for auditory perception. Patients with auditory dys-synchrony have difficulty in perceiving timing-related information, but not intensity or frequency related information. Individuals with auditory dys-synchrony seem to have difficulty in - discriminating between low frequencies, temporal integration, gap detection, modulation detection, perceiving sounds in the contexts of preceding or succeeding sounds, detecting beats for binaurally presented stimuli and using interaural time cue to localize sound. The nature of these deficits in patients with auditory dys-synchrony is different from those seen in individuals with cochlear hearing loss.

3.4 Speech Perception

Speech perception abilities in individuals with auditory dys-synchrony are related to temporal processing abilities (Ranee et al., 2004; Zeng et al.,

1999; Zeng, et al., 2005). Subjects in whom cortical evoked potentials could be recorded generally showed higher speech perception scores (Ranee et al., 2002). Rance et al. showed that cortical evoked potentials could be elicited using both tonal and speech stimuli in children with auditory dys-synchrony. Presence of cortical evoked potentials with age appropriate latency and morphology seemed to be related to open set speech identification scores and benefit derived by subjects from amplification.

Kraus et al. (2000) systematically examined fine-grained speech perception abilities in an adult with auditory dys-synchrony. They measured just noticeable differences (JNDs) for three CV continua: /ba-wa/, /da-ga/ and another /da-ga/ continuum in which amplitude of formant transitions had been enhanced. Results showed that on /ba-wa/ continuum, the JND of the patient with auditory dys-synchrony was comparable to that of normal hearing subjects. However, JND for /da-ga/ continuum in auditory dys-synchronics was poorer than in normal hearing adults. Normal hearing adults could discriminate between stimuli in which the onset frequency of the 3^ formant frequency differed by 80 Hz in the two stimuli, while individual with auditory dys-synchrony required a difference of almost 120 Hz between the stimuli. When the amplitude of formant transition was enhanced in /da-ga/ continuum relative to vowel segment, the subject with auditory dys-synchrony showed even greater difficulty in discriminating along the continuum. These data suggest that the subject was able to discriminate between stimuli in which the transition was slow, but had difficulty in discriminating stimuli that were characterized by rapid spectro-temporal changes throughout the formant

transition. Enhancement of intensity related information in the formant frequencies did not improve speech perception in the subject with auditory dys-synchrony. In the same subject, Kraus et al. (2000) measured word identification by manipulating three factors, namely, signal to noise ratio, lexical difficulty, and number of talkers (single vs. multiple talkers) to investigate the effect of multiple sources of variability and signal degradation on speech perception. The subject with auditory dys-synchrony showed marked effect of noise on speech identification compared to normal hearing subjects. However, on tasks where the number of talkers and lexical difficulty was varied, the performance of patient with auditory dys-synchrony was similar to that of normal hearing subjects. On the same subject, Kraus et al. demonstrated the presence of M M N for /ba-wa/ contrast and its absence for the /da-ga/ contrast. This result shows that individuals with auditory dys-synchrony have problem in perceiving dynamic cues at the onset of the stimulus. These results are consistent with findings from behavioural studies which showed normal performance in the discrimination of /ba-wa/ contrast, but poor performance on the /da-ga/ continuum.

The data of Kraus et al. (2000) amply demonstrate that individuals with auditory dys-synchrony perform similar to normals on speech perception if the rate of change of temporal features is slow (as in /ba-wa/ continuum), but their speech processing mechanism breaks down while dealing with rapid spectro-temporal changes. As many speech events critical for the perception of stop consonants involve rapid temporal changes at the onset of the stimulus (for example burst and transition), it is logical to predict that auditory dys-

synchronies will have problems in speech perception. Presence of normal cortical evoked potentials in some auditory dys-synchronics suggests that the auditory cortex does function normally even if signal encoding is not so normal at the lower levels. Redundancy provided by the system at various levels facilitates this. This being the case, it follows that if signal coding /processing can be improved at lower levels in the auditory pathway (for example, the brainstem), then signal perception will all the more be better at cortical level. Perhaps, this is the challenge before us in designing hearing aids or in developing algorithms for digital signal processing.

4.0 Management

Management of persons with auditory dys-synchrony is a challenge for audiologists because of the inherent heterogeneity of the population in terms of audiological and neurological findings, etiology and pathophysiology. For reasons that are not yet known, 7 to 10% of the patients with auditory dys-synchrony show no observable symptoms other than absence of ABR. Some of them will also develop normal speech and language with only complaints of difficulty in understanding speech in noise. Generally, hearing aids are not of much benefit for individuals with auditory dys-synchrony. Ranee, Con-Wesson, Wunderlich and Dowell (2002) found that 50% of their patients in whom they could elicit cortical evoked potentials also benefited from hearing

aids. However, this could be an over estimation because these subjects had not been followed up for any length of time to warrant such a conclusion on the benefit from hearing aids. Berlin et al. (2003) reported that none of the 200 individuals (in their database) with auditory dys-synchrony had successfully adapted to hearing aids or oral language as life long communication strategy.

There is increasing evidence that children with auditory dys-synchrony benefit from cochlear implantation. Peterson et al. (2003) compared two groups of ten children each - one group with auditory dys-synchrony and cochlear implantation and the other with cochlear hearing loss and cochlear implantation. These two groups of children were compared for aided and unaided audiograms, performance on age appropriate speech perception tests, parental report of benefit from cochlear implant, electrically evoked ABRs and visually detectable electrical stapedial reflex. The results showed no important difference between the two groups with respect to benefits from cochlear implant. Peterson et al. (2003) concluded that the use of cochlear implants is a viable option for selected children with auditory dys-synchrony. Similar benefits of cochlear implantation in children with auditory dyssynchrony have also been reported by other investigators (Buss, Labadie, Brown et al., 2002; Miyamoto, Kirk, Renshaw, & Hussain, 1999; Shallop, Peterson, Facer, Fabry, & Driscoll, 2001; Trautwein, Sininger, & Nelson, 2000). It has been shown in animal studies that electrical stimulation produces synchronous ABRs even when the peripheral auditory nerve is demyelinated (Zho, Abbas & Assouline, 1995). This suggests that improved speech

perception in auditory dys-synchronic children fitted with cochlear implants could be the result of better synchrony in neural firing in response to electrical stimulation. In addition, the electrical stimulation provided by cochlear implants may even promote neural survival by preventing deprivation.

It is recognized, however, that the degree of hearing loss may not justify the use of cochlear implant in some of these individuals with auditory dys-synchrony as an option of management. It is also right to say that the high cost of cochlear implantation deters a majority of individuals with auditory dys-synchrony, particularly in developing countries like India, from choosing cochlear implant as an option of management. Therefore, though hearing aids may not bring in the expected benefits as they do not address issues relating to processing of temporal information in speech, they still seem to be a better option of management because of their lower cost. However, there is need for algorithms in hearing aids which enable optimal retention of temporal information of speech while processing. This, however, points to the need for understanding ways in which temporal information is degraded/distorted while they are processed in hearing aids. Research on identification of temporal features of speech that are important for normal or near normal speech perception in individuals with auditory dys-synchrony is also warranted.

5.0 Importance of Temporal Cues in Speech Perception

Aspects of time is of essence in the perception of a serial event like speech. Near perfect speech recognition can be achieved with minimum spectral information as long as temporal cues are available (Shanon, Zeng, Kamath, Wygonski & Ekelid 1995; Turner, Souza & Forget, 1995). Shanon et al. (1995) showed that even under the condition of reduced spectral cues, slowly varying temporal fluctuations (<50 Hz) can result in relatively high speech recognition in normal hearing subjects. Turneret et al. (1995) showed that subjects with cochlear hearing loss can effectively use temporal cues in speech. Different techniques have been employed to improve speech perception abilities in both hearing impaired and normals by enhancing temporal cues of the speech. Some of these strategies are described in the following section.

5.0.1 Transition Enhancement

Formant transitions have been shown to be of importance in differentiating place of articulation in stop consonants. These rapid transitions occur in the beginning of stop consonants and are brief. It has been shown that children with dyslexia, learning problems, specific language impairment and elderly listeners have difficulty in perceiving stop consonants due to poor processing of dynamic cues (Tallal & Piercy, 1975; Tallal et al., 1996;

Meizenich et al., 1996; Strouse, Ashmead, Ohde, & Grantham, 1998). One of the techniques that has been advocated to overcome poor temporal processing abilities such as those observed in these clinical populations is time scale modification of the speech. Tallal and Piercy (1975) investigated the effect of lengthening of formant transition on syllable discrimination in normals and dysphasics. They found that syllable discrimination abilities of dysphasics were similar to those of normals when formant transitions were lengthened, but were poorer than those of normals at shorter transitions.

Keatiny and Blumstein (1978) evaluated the effect of transition length on the perception of /ba/ and /ga/ in normal adults. Lengthening of formant transition resulted in better "within-the-category" discrimination indicating thereby that more auditory information is available at longer transition lengths. Modification of consonant duration - CV amplitude ratio improves speech intelligibility even under adverse listening conditions in both young and elderly listeners. But, modification of consonant duration alone does not significantly improve speech recognition. However, increasing consonantal duration resulted in a reduction in the frequency of consonant confusions (Gordon-Salant, 1986).

In their seminal work, Tallal et al. (1996) and Meizenich et at. (1996) demonstrated that training with temporally modified speech resulted in significant improvement in speech discrimination as well as language comprehension in a group of 11 children with language learning impairment compared to age and gender matched language impaired children who were

trained with natural, unmodified speech. Acoustic modifications of speech carried out by Tallal et al. (1996) included: (a) time scale modification (doubling the duration) without altering spectral content, and (b) differential enhancement of fast modulations in the range of 3 to 30 Hz. These modifications were based on the previous observation of Tallal and Piercy (1975) that prolongation of formant transitions increased intelligibility of speech in the language learning impaired. Children with language impairment received training for 4-weeks. A comparison of pre - post training test performance showed significant improvement in speech discrimination and language proficiency following training. These improvements were significantly more in the experimental group compared to the control group which was trained with precisely the same exercises, but, with natural unmodified speech. Based on these results, Tallal and her colleagues developed a computer-based language intervention programme called Fast Forword® for children with language learning impairment, auditory processing deficits, dyslexics etc. Field studies on more than 500 children identified to have language learning impairment have shown significant improvements in the receptive and expressive language skills of these children when trained with Fast Forword® (Tallal & Merzenich, 1997). However, Lacerda (2001) reported that varying length of transition does not affect identification (labelling) of stimuli. But, discrimination of CV contrast was marginally better for lengthened transition CVs. Sundberg and Lacerda (2003) reported significant improvements in the language comprehension of severely language-impaired children following training with temporally

modified (lengthened) or spectrally modified stimuli (enhanced amplitude of formants and burst). Tempel et al. (2003), in an fMRI study, provided evidence for improvement in neural mechanisms in dyslexic children who underwent a behavioural remediation program with acoustically modified speech. Physiologically, these children showed increased activity in multiple areas in the brain. There was a strong correlation between increased neural activation in the left temporo-parietal cortex and improvement in oral language abilities.

Bradlow, Kraus, Nicol, McGee and Cunnigham (1999) investigated the effect of lengthened formant transition in a group of children with learning problems. They studied discrimination thresholds along two separate /da-ga/ continua. The continua differed from each other only in the duration of formant transitions. Results showed that lengthening of transition did not improve speech discrimination, but resulted in significant improvement in MMN responses. They concluded that lengthening transition duration could result in enhanced encoding of the signal at pre-attentive neural level. However, this manipulation was not sufficient, on its own, to facilitate perceptual discrimination in children with learning problems.

Thus there is evidence from behavioural (Tallal et al., 1996; Meizenich, et al., 1996), physiological (Tempel et al., 2003) and electrophysiological (Bradlowet, et al., 1999) studies to show that lengthening of fbrmant transition augments speech perception skills in some clinical populations like language learning impaired and dyslexic.

5.0.2 Enhancement of Temporal Envelope

Expanding the temporal envelop of speech has been shown to result in an improvement in speech recognition in noise and ease of listening in both normal and hearing impaired (Apoux, Crouzet & Lorrenzi, 2001; Lorrenzi, Berthommier, Apoux & Bacri, 1999). These investigators extracted low frequency temporal modulations and raised it to the power two. The resulting envelopes were then used to modulate white noise. This resulted in the degradation of spectral information and forced the listeners to identify speech using temporal (envelope) cues primarily. The expansion of temporal envelope resulted in significant improvement in the ease of listening as well as in speech recognition scores under noisy conditions. Expansion of temporal envelope at low modulation frequencies (< 16 Hz) improved speech identification in noise in both normal and hearing impaired listeners (Apoux, Tribut, Dehruille & Lorenizie, 2004).

5.0.3 Clear speech

Sentences spoken 'clearly' are significantly more intelligible than those spoken 'conversationally' for hearing- impaired listeners in a variety of backgrounds (Picheny, Durlach, & Braida, 1985, 1986; Uchansky, Choi, Braida, Reed, & Durlach, 1996). Acoustical properties of clear speech are: (a)

reduced speaking rate which is achieved by lengthening individual speech sounds or by inserting pauses in between, (b) increase in spectral energy at high frequencies, (c) increased modulation depth, and (d) lengthening of voice onset time of voiceless stop consonants (Kruase & Braida, 2004; Picheny, et al., 1985, 1986; Uchansky et al., 1996). Evidence for better encoding of clear speech than conversational speech has also come from animal studies. Cunningham, Nicol, King, Zecker and Kraus (2002) recorded aggregate neural responses from auditory midbrain, thalamus and auditory cortex of anesthetized guinea pigs to a syllable /ada/ in quiet and in noise. Neural encoding for the steady state vowel portion of the syllable was more resistant to degradation than neural encoding of the dynamic portion (release burst and transition) of the consonant. Furthermore, when the consonant was enhanced to simulate clear speech by increasing the duration of stop gap and intensity of release burst, it resulted in better representation of the consonants at auditory midbrain, thalamus and auditory cortex.

5.2. Temporal Processing and Speech Perception

The question of whether or not deficits in temporal processing abilities (measured using non-speech stimuli) lead to impaired speech perception has not been answered unequivocally. Errors made by hearing-impaired listeners in their speech perception have been attributed to reduced or deviant

processing of temporal information (Erber, 1972; Gordon- Salant & Fitzgibbons, 1993; Price & Simon, 1984; Tyler, Summerfield, Wood & Fernandes, 1982). Many studies have shown a strong correlation between listeners' ability to perceive short duration acoustic signals and speech perception. Gap detection abilities are linked to listener's ability to process time events related to distinctions in speech between voiced and unvoiced cognates, and various manners of syllable transition (DeFillippo & Snell, 1986). Dreschler and Plomp (1985) showed that impaired speech perception is related to both frequency and temporal resolution. Strouse, et al. (1998) reported marked deficits in temporal processing abilities such as gap detection threshold, intra aural time differences and voice onset time perception in elderly listeners with normal hearing. However, they failed to show any relation between perception of VOT and other psychophysical tasks. Several investigators have observed a significant correlation between gap detection thresholds and speech recognition scores in the hearing impaired even when audiometric thresholds were factored out. Speech perception in noise and reverberation is strongly related to gap detection thresholds (Glassberg & Moore, 1989; Tyler, et al., 1982; Gordon- Salant & Fitzgibbons, 1993; Snell, Mapes, Hickman, Frisina, 2002). Taken together, these results suggest that poor speech perception might be partly related to difficulties in processing temporal features.

Thus, it is conceivable that perceptual problems with short duration stimuli such as gap detection and poor speech perception abilities in individuals with hearing impairment emanate from the same abnormality of

processing the information nature of which is yet to be identified. It is widely accepted that many speech events that are critical for speech perception are short in duration. By making these cues more prominent, speech perception abilities of children with learning impairment, elderly, and dyslexic can be enhanced. We have seen from the above review that individuals with auditory dys-synchrony have difficulty in processing rapid and brief acoustic events. Retaining these temporal events as distortionless as possible after digital signal processing in hearing aids may be the strategy to make hearing aids more beneficial to individuals with auditory dys-synchrony.

The review in the paragraphs above throws up areas where there are gaps in information, particularly in the Indian context. There is absolutely no information on the prevalence of the problem of auditory dys-synchrony in India. Results of Sininger and Oba (2001) suggest that the onset of the problem takes place very early in life - during the language acquisition years. It is absolutely imperative that such children are identified early, and input crucial to development of their auditory and perceptual skills provided.

Another important area where we require more information is the speech perception abilities, specifically processing of temporal information, in persons with auditory dys-synchrony. There is reasonable evidence to say that auditory dys-synchronics have difficulty in processing temporal information. But, the question is on the specific timing-related parameters that are affected and the specific ways they are affected. Related to this is also the question of the specific ways in which temporal parameters are to be modified to enhance

speech perception in auditory dys-synchronics. This issue assumes importance because there are conflicting reports about the usefulness of amplification (hearing aids) for auditory dys-synchronics. Though there are reports about the positive benefit of cochlear implants to auditory dys-synchronics, the cost of the implant and the degree of hearing loss restrict the choice of cochlear implants as a management strategy. Therefore, there is no other solution than to improve speech processing, particularly those relating to timing, in hearing aids to make them friendly to auditory dys-synchronics. The present study is highly justified in this context.

CHAPTER 3

METHOD

The objectives of this study were to (a) estimate the prevalence of auditory dys-synchrony in the South Indian city of Mysore, (b) evaluate perception of temporal parameters of speech in individuals with auditory dys-synchrony, and (c) measure modulation detection thresholds for sinusoidally amplitude modulated white noise in individuals with auditory dys-synchrony. In the first phase, prevalence of auditory dys-synchrony was estimated through a retrospective register-based study. Perception of three temporal parameters of speech namely, transition duration (TD), burst duration (BD) and voice onset time (VOT) were investigated for their influence on speech perception in individuals with auditory dys-synchrony in the second phase of the study. Four experiments were designed to achieve the objectives of the second phase. Speech identification scores for unmodified stimuli were measured in Experiment 1. The second experiment was on determining just noticeable difference (JND) for TD, BD and VOT in CV (consonant + vowel) syllables. The effect of lengthening of each of these temporal parameters on speech perception in individuals with auditory dys-synchrony was evaluated in the third experiment. The fourth experiment was designed to evaluate perception of speech in individuals with auditory dys-synchrony for CV stimuli which had JNDs which had yielded the maximum speech identification score in Experiment 3. Modulation detection thresholds for sinusoidally amplitude

modulated white noise at 4 Hz, 16 Hz, 32 Hz, 64 Hz, 128 Hz and 200 Hz were measured in Phase III of the study.

1.0 Phase I: Prevalence of the Problem

A register based study design was employed to find the prevalence and to describe the audiological characteristics of persons with auditory dys-synchrony. A retrospective analysis of case reports was done wherein test findings of all those who visited the Department of Audiology, All India Institute of Speech and Hearing, Mysore between January 2000 and December 2003 were reviewed. Mysore is a major city in South India and has a population, predominantly rural, of about a million. People in this part of the country speak Kannada - a Dravidian language (Jayaram, 1984). A total of 21,236 (11,712 males and 9524 females) records were reviewed. 11,205 of these were of persons with permanent sensori-neural hearing loss. 61 of these 11,205 (5854 males and 5351 females) individuals were identified as cases of auditory dys-synchrony. Criteria employed to identify auditory dys-synchrony were those recommended by Starr, Sininger & Praat (2000). Accordingly, patients who were identified to have auditory dys-synchrony

a) had preserved cochlear amplification, that is, presence of transient evoked otoacoustic emissions,

b) showed altered auditory nerve responses (absent, but, if present then abnormal ABRs),

c) showed no evidence of space occupying lesion on neurological examination, and

d) had a normal otological and tympanometric study.

All these individuals had undergone (a) puretone, speech and immittance audiometry, and (b) auditory brainstem response and otoacoustic emissions evaluation. However, otoacoustic emission testing had been done only when ABR was abnormal. Roughly, 20% of the patients had undergone neurological examination, but all the 61 individuals who were eventually diagnosed to have auditory dys-synchrony had been subjected to all tests including neurological investigation. Neurological evaluation included a clinical examination, an MRI or a CT scan.

1.1 Testing procedure and instruments

It was ascertained from case records that all these subjects had been tested under standard conditions. All subjects had been tested with calibrated (ISO, 389) audiometers in sound treated rooms. Puretone testing had been done using modified version of Hughson and Westlake procedure. Speech identification testing had been done with live voice presentation of phonetically balanced monosyllables at maximum comfortable level. Immittance evaluation (tympanometry and acoustic reflex threshold testing) for 226 Hz probe tone had been carried out with calibrated middle ear analysers (GSI-33 or Tympstar). Auditory brainstem response testing had been done using either Biologic Navigator or Nicolet Bravo evoke potential system. Identical protocol had been employed with all the patients. It was checked from records that auditory

brainstem testing had been done twice to ensure reproducibility of waveforms. A group of 30 normal hearing adults had been tested to establish the nHL values. Results showed that 0 dB nHL click had a peak equivalent SPL of 26 dB. Transient evoked otoacoustic emissions (ILO 292) had been measured in a sound treated room for clicks at 80 dB ± 5 dB peSPL. An emission had been considered to be present if the waveform reproducibility was more than 50% and the overall signal to noise ratio was more than 3 dB at two frequency bands at least.

2.0 Phase II: Speech perception

2.1 Subjects

Two groups of subjects, namely, a group of persons with confirmed diagnosis of auditory dys-synchrony, and an age and gender matched group of normal hearing subjects participated in this phase.

2.2 Auditory Dys-synchrony Group

30 subjects, in the age rage of 16 to 30 years (mean age 22.4 years), with a confirmed diagnosis of auditory dys-synchrony, formed the auditory dys-synchrony group (hereinafter referred to as Group A). This age range of 16-30 years was selected as it has been reported that psychoacoustical abilities reach a

plateau in this age range (Lynne, Werner & Gray 1998). Only those subjects who

* exhibited signs and symptoms of difficulty in understanding speech,
* showed no symptoms of external or middle ear problems,
* showed no history of ototoxic drug usage or exposure to loud noise,
* had not undergone any formal training in auditory learning activities, and
* whose primary language (mother tongue) was Kannada were selected for the study.

This information was elicited through a structured clinical interview and case history to begin with. Furthermore, subjects in the auditory dys-synchrony group had to manifest the following audiological characteristics to be included in the study:

* Pure tone hearing thresholds less than 70 dB HL (ISO 389) at lOOOHz, 2000Hz, and 4000Hz with no air-bone gap.
* Disproportionate speech identification scores in relation to puretone hearing loss. Expected speech identification scores were based on Vanaja and Jayaram (2003).
* "A" type tympanogram with no ipsilateral and contralateral reflexes.
* No ABRs, but, if present, severely abnormal auditory brainstem responses at 90 dB nHL.

* TEOAEs with normal or robust amplitude with the magnitude of contralateral suppression of evoked otoacoustic emission less than 0.5 dB.

* Abnormal masking level difference.

Air conduction and bone conduction puretone thresholds were determined using a calibrated clinical audiometer (Maico MA 53, calibrated as per ISO 389) in a sound treated room with ambient noise at permissible limits for audiometry (ANSI, 1991). Audiometer was calibrated in the beginning of the experiment and every two months thereafter using a sound level meter (Quest 1800) and a pressure microphone (Quest 4170). Puretone audiometry was carried out using modified version of Hughson and Westlake procedure using 5 dB step size.

Vandana's (1998) Speech Identification Test in Kannada was used to assess open set speech perception abilities in these subjects. This test consists of 50 bisyllabic meaningful words of Kannada. Validity and reliability of this test has been established on native speakers of Kannada (Vandana, 1998).

Tympanometry and reflexometry was done using a calibrated immittance meter (GSI- Tymp star V 2, calibrated as per ANSI, 1987). Tympanograms were obtained for 226 Hz probe tone. Ipsilateral and contralateral acoustic reflex thresholds were measured at 500 Hz, 1000 Hz, 2000 Hz and 4000 Hz.

Table 3.1 shows the protocol used for auditory brainstem response testing (Nicolet Bravo). Waveform was recorded twice to ensure reproducibility of recordings. A group of 30 normal hearing adults were tested

to establish nHL values. Results showed that 0 dB nHL click had a peak equivalent SPL of 26 dB.

Table 3.1. Standard protocol used for ABR testing in this study.

Stimulus parameters		Acquisition parameters	
Stimulus	Clicks	Filter	30-3000 Hz
Polarity	Rarefaction	Montage	Cz-Al and Cz-A2
Level	90dBnHL	Window	15 ms
Duration	100 us	Artifact rejection	>50uV
Transducer	Electrically shielded head phones		
Number of sweeps	2000		
Rate	ll.l/s		

Transient evoked otoacoustic emissions (ILO 292) were measured in a sound treated room for clicks at 80 dB + 5 dB peSPL. An emission was considered to be present if the waveform reproducibility was more than 50% and the overall signal to noise ratio was more than 3 dB, at least at two frequency bands. OAEs were recorded again in the presence of contralateral broadband noise at 30 dB SL (re: threshold noise), presented through insert receiver of a calibrated clinical audiometer (Maico, MA- 52) to measure the magnitude of contralateral suppression of otoacoustic esmissions. Care was taken to ensure that the position of the probe was not altered during the two

recordings. Suppression magnitude of more than 0.5 dB was considered normal.

Masking level difference (MLD) was done at 250 Hz and 500 Hz. MLD was calculated as the difference in threshold between SoNo condition and SnNo condition. A difference of more than 10 dB between the two conditions was considered normal.

2.3 ENT and Neurological Evaluation

All the subjects, before being selected for the study underwent an ENT examination which was done to rule out any external or middle ear problem. Similarly, all subjects were subjected to a neurological examination by a qualified neurologist for ruling out any peripheral neuropathy or space-occupying lesion. Neurological evaluation also included a CT or an MRI, if required.

2.4 Normal Subjects

30 normal subjects, with normal hearing, and matched for age and gender, constituted the normal group. The normal subjects participated only in Phase II and III of the study. All the normal subjects were native speakers of Kannada. It was ascertained from a structured interview that none of the normal hearing subjects selected for the study had difficulty in understanding speech in daily listening conditions, and that they did not have any history of neurologic

or otologic disorder. Subjects to be included in the normal group had to show the following audiological findings:

> Puretone thresholds < 15 dB HL (ISO, 389) at octave frequencies between 250 Hz to 8 kHz.

> 95% to 100% speech identification scores at 40 dB SL (ref: average hearing thresholds at 500 Hz, 1000 Hz and 2000 Hz).

> "A" type tympanogram with ipsilateral and contralateral reflexes at normal sensation levels.

> Identifiable auditory brainstem response peaks (wave I, III and V)

> Normal amplitude TEOAEs with magnitude of contralateral suppression greater than 0.5 dB.

> Masking level difference of more than 10 dB.

> Normal results on dichotic CV (Krishna, 2002) and gap detection test (Shivaprakash, 2003). Dichotic CV test consisted of 30 dichotically paired CV stimuli recorded on a compact disc. Standardization of this test on adults has already been done (Krishna, 2002). Gap detection test measures the listener's ability to detect silences embedded within bursts of noise. 3-interval alternate forced choice method was employed to measure gap detection thresholds. Results were considered to be normal if the values fell within mean ± 1 SD of normative data (Shivaprakash, 2003).

Dichotic CV and gap detection tests were administered only on normal subjects to rule out central auditory disorder. These tests were not administered

on auditory dys-synchrony group as the results of these tests are likely to be affected in the presence of peripheral (auditory nerve) impairment. Instrumentation, test environment and administration of all audiological tests was the same for both groups of subjects.

2.5 Parameters Tested

Temporal parameters of speech that were considered in this study were transition duration, burst duration and voice onset time. Transition duration is the interval during which there are rapid spectral changes in the fbrmant frequencies after the release of occlusion. Burst duration is the interval between the onset of the burst and the release of articulators. Transition duration and burst duration are those dynamic spectral cues, at the onset of stimulus, that help in perceiving the place of articulation for stop consonants. Psychophysical experiments have shown that perception of dynamic acoustic information is affected, especially at stimulus onset, in patients with auditory dys-synchrony (Kraus et al., 2000; Starr, et al., 1991). Improvement in speech perception has been reported for stimuli with lengthened transitions in clinical groups like dyslexics, language learning impaired children and dysphasics after training (Tallal, et al., 1996; Meizenich et al., 1996).

V O T is a strong perceptual cue for voicing in most languages and is language dependent. V O T is the interval between the release of the stop and the onset of voicing in unvoiced consonants of Kannada while in voiced consonants, it is the duration of voicing pulses preceding the burst. V O T was

considered in this study because psychophysical experiments using non-speech stimuli have shown that individuals with auditory dys-synchrony have difficulty in processing short duration sounds, but not long duration sounds (Zeng, Oba, Garde, Sininger, & Starr, 1999; Zeng, Kong, Michalewski, & Starr, 2005). VOT is a short duration temporal cue that helps in perceiving voicing information.

2.6 Test Stimuli

CV syllables with voiceless stop consonants - velar /k/, alveolar /t/, retroflex /t/, and bilabial /p/ - and their voiced cognates were used in the study. These consonants were paired with vowel /a/ to get CV syllables. A 25-year-old male native speaker of Kannada uttered these syllables, one after the other. The spoken syllables were digitally recorded on a data acquisition system with a 32-bit analog to digital converter, and at a sampling frequency of 44.1 kHz. The syllables thus recorded were edited using Praat software. TD, BD and VOT were identified on both waveform and spectrogram of the CV syllables. Each of these parameter was lengthened in 5 ms step by means of Pitch Synchronized Overlap and Add technique (PSOLA). PSOLA allows lengthening and shortening of the stimulus in the time domain without affecting the physical characteristics of the stimulus such as spectral shape, amplitude distribution, and periodicity (Moulines & Laroche, 1995).

2.7 Procedure

2.7.1 Experiment 1: Speech Identification Scores with Unmodified Stimuli

This experiment was carried out on subjects in both the auditory dys-synchrony and normal hearing group. Ten repetitions of each of the eight unmodified experimental stimuli were randomly presented and speech identification scores for the original unaltered stimuli were noted down. Stimuli were played through a personal computer at a sampling frequency of 44.1 kHz which was later inputted into a calibrated clinical audiometer (Maico MA-53). Audiometer was calibrated in the beginning of the experiment and every two months thereafter (Quest 1800 sound level meter and Quest 4180 free field microphone). Subjects received the stimuli through a loudspeaker connected to the audiometer. Loudspeaker was positioned at a distance of I meter and at 90° azimuth. Presentation level was kept constant at 40 dB SL (ref: average threshold at 500 Hz, 1 kHz and 2 kHz) for all the subjects. Written responses were obtained from the subjects, if they are literate. If subjects were illiterate, experimenter and another native speaker of Kannada noted down the responses of the subjects independently, and only those responses which had 100% agreement between two observers were taken for the further analysis.

2.7.2 Experiment 2: JNDs for Temporal Parameters of Speech

Subjects in both the auditory dys-synchrony and the normal hearing groups participated in this experiment. JNDs for the three different temporal parameters - transition duration, burst duration and voice onset time - were established in this experiment. JND was determined using an adaptive tracking technique - Parameter Estimation through Sequential Testing (PEST) with an "AX same-difference" discrimination paradigm. 'A' is anchor stimulus and 'X' is variable stimulus in this paradigm where the subjects' task is to indicate whether 'A' is the same as 'X' or not. PEST is an adaptive procedure that considers changes in both direction and step size of the stimulus to reach the target level. Each of the original stimuli, called anchor, was paired with another stimulus designated as the variable stimulus. The variable stimulus, as the name suggests, is variable. Variable stimuli were those CV syllables in which the TD, BD and VOT, as the case may be, had been lengthened. Figures 3.1-3.4 show a representative waveform and spectrogram of the unmodified and temporally modified stimulus /ba/. Figure 3.5 shows the spectrum of unmodified and temporally modified stimulus. It may be noted that temporal modification through lengthening of TD, BD and VOT did not alter the spectrum of the sound.

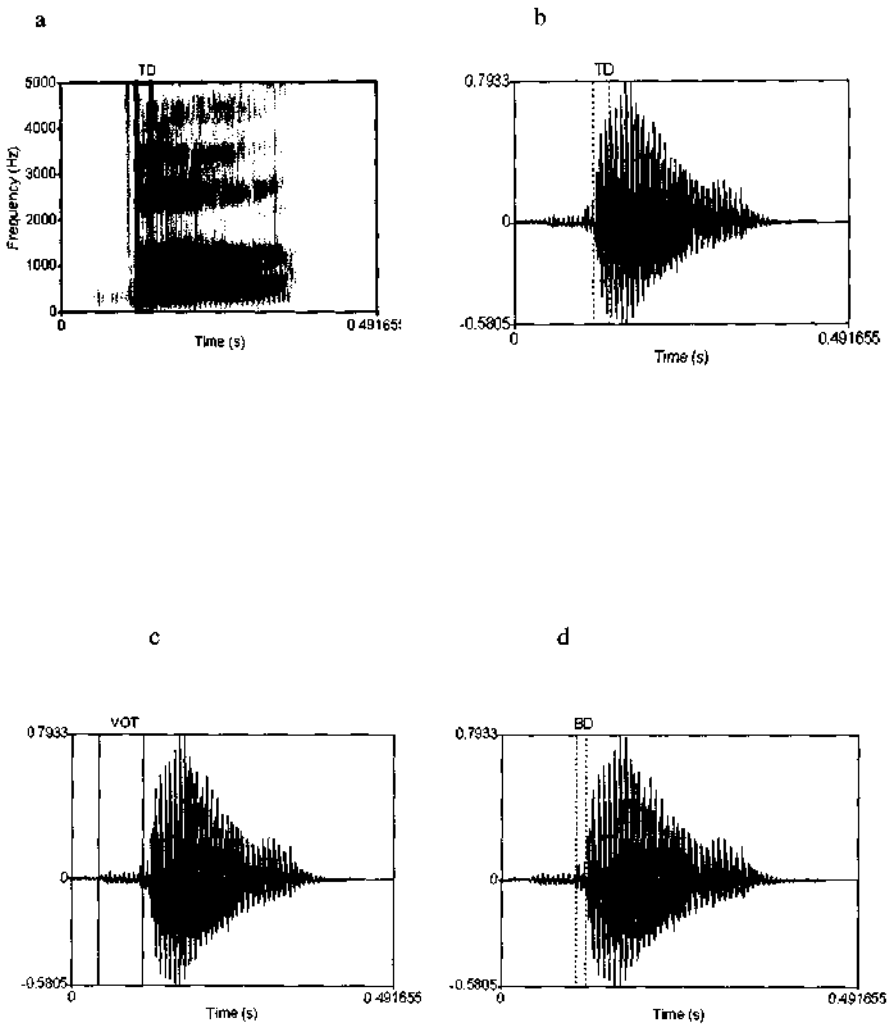


Figure 3.1. a) Spectrogram of unmodified stimulus /ba/. (b) Waveform of unmodified stimulus /ba/. TD is marked with a rectangle, (c) V O T of the unmodified stimulus /ba/ (d) BD of the unmodified stimulus /ba/

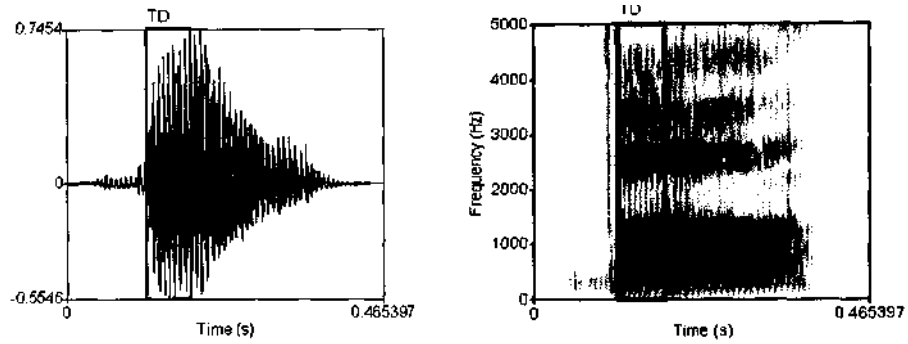


Figure 3.2.Waveform and spectrogram of stimulus /ba/ whose transition duration was lengthened by 30 ms.

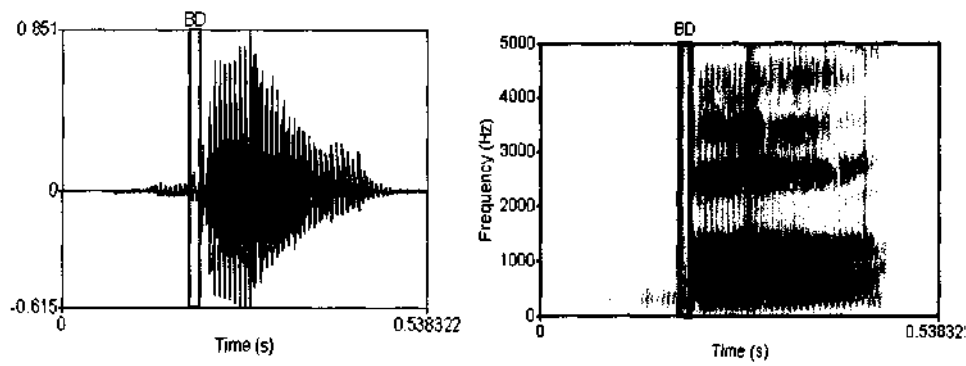


Figure 3.3. Waveform and spectrogram of stimulus /ba/ whose burst duration was lengthened by 10 ms.

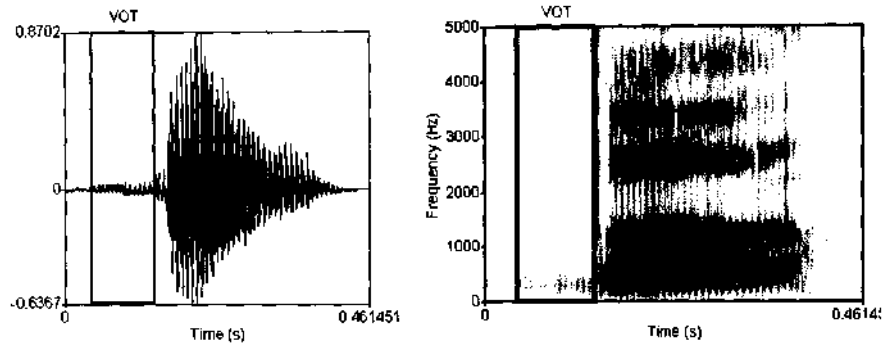


Figure 3.4. Waveform and spectrogram of stimulus /ba/ whose voice onset time was lengthened by 10 ms

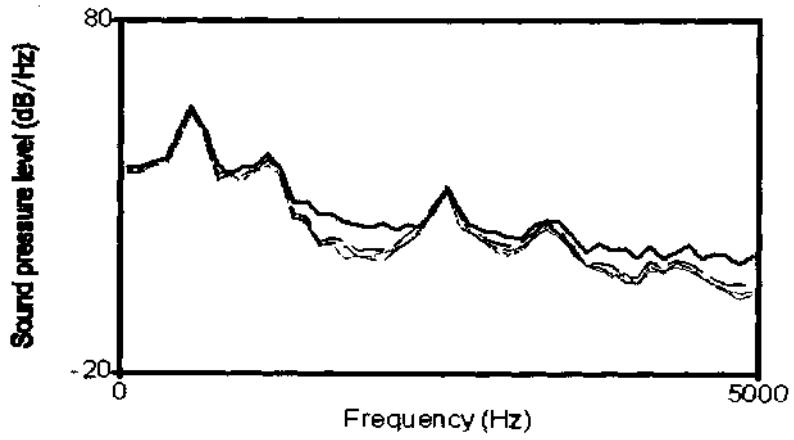


Figure 3.5. Spectrum of the unmodified (black dashed line), transition duration lengthened by 30ms (red dashed line), burst duration lengthened by 10ms (black solid line) and voice onset time lengthened by 30 ms (blue dashed line) stimulus /ba/. Note that all the modification did not bring any significant change in the spectrum

Inter stimulus interval between the anchor and the variable stimulus was 500ms. Step size and the direction of variable stimulus was changed according to rules of PEST (Taylor & Creelman, 1967). Minimum step size was 5 ms. Instrumentation, test environment and presentation level was the same as in Experiment 1. IND for voice onset time for /pa/ was determined as shown in Table 2, for example.

Table 3.2. An example of sequence of presentation of stimuli in Experiment 2

/pa/ with VOT 20ms	Anchor stimuli (A)	Variable stimuli (X) (lengthened with PSOLA)	Response
First pair	/pa/ with VOT 20ms	/pa/ with VOT 60 ms*	different
Second pair	/pa/ with VOT 20ms	/pa/ with VOT 50 ms	different
Third pair	/pa/ with VOT 20ms	/pa/ with VOT 40 ms	different
Fourth pair	/pa/ with VOT 20ms	/pa/ with VOT 20 ms	same
Fifth pair	/pa/ with VOT 20ms	/pa/ with VOT 30 ms	different
Sixth pair	/pa/ with VOT 20ms	/pa/ with VOT 25 ms**	same
Seventh pair	/pa/ with VOT 20ms	/pa/ with VOT 30 ms	different
Eight pair	/pa/with VOT 20ms	/pa/ with VOT 25 ms	same
Ninth pair	/pa/ with VOT 20ms	/pa/ with VOT 30 ms ms	different
Tenth pair	/pa/ with VOT 20ms	/pa/ with VOT 25 ms ms	same

*only an example

** minimum step size was reached.

2.7.3 Experiment 3: Speech Identification Testing with Temporally Modified Stimuli

Only subjects in the auditory dys-synchrony group participated in the third experiment. In this experiment, each of the selected temporal parameters was modified (lengthened) in terms of JNDs and the effect of modified stimulus on speech identification was determined.

For example, to evaluate the effect of changes in VOT on the perception of /pa/, four tokens were generated. First, a /pa/ with a VOT of one JND for normal listeners; second, a /pa/ with a VOT of two JNDs for normal listeners; third, a /pa/ with a VOT of three JNDs for normal listeners, and fourth, a /pa/ with a VOT of four JNDs for normal listeners. Otherwise, the procedure of the experiment was similar to the one described in Experiment 1. Ten repetitions of each token were presented. The presentation stopped at the level where the subject identified the correct sound or the sound which elicited the closest response as determined by a feature analysis. The purpose of this experiment was to find out the characteristics of the stimulus with modified temporal parameters, which resulted in the best identification of speech segments in persons with auditory dys-synchrony. Of these tokens, the one which was correctly identified as target stimulus (/pa/ in this example), or the one which elicited the closest response to the target sound, as determined by a feature analysis, will be the one to be achieved in hearing aids for patients with auditory dys-synchrony. Feature analysis to determine the closest response to the target sound considered voice and place of articulation information. The data gathered

in this experiment was utilized in Experiment 4. All the stimuli were presented in random order. As the stimuli were isolated syllables and were presented in random order, learning and practice effect can be discounted. The JND values of subjects with auditory dys-synchrony are not considered in the tokens generated for this experiment because a pilot study showed a high heterogeneity in the responses of the auditory dys-synchrony group.

2.7.4 Experiment 4: Speech Identification Testing with Stimuli with the Optimal Values of JNDs

A synthetic token was generated for each of the eight CV syllables wherein the TD, BD and VOT corresponded to the level which resulted in the best identification of the target stimulus in Experiment 3. The purpose of Experiment 4 was to determine speech identification scores for stimuli which had the best values of TD, BD and VOT as determined in Experiment 3. The procedure, test environment, instrumentation and presentation level were the same as in Experiment 1. These scores were compared with the speech identification scores obtained for the original unaltered stimuli. Figure 3.6 is a block diagram of the sequence of experiments in Phase II of the study.

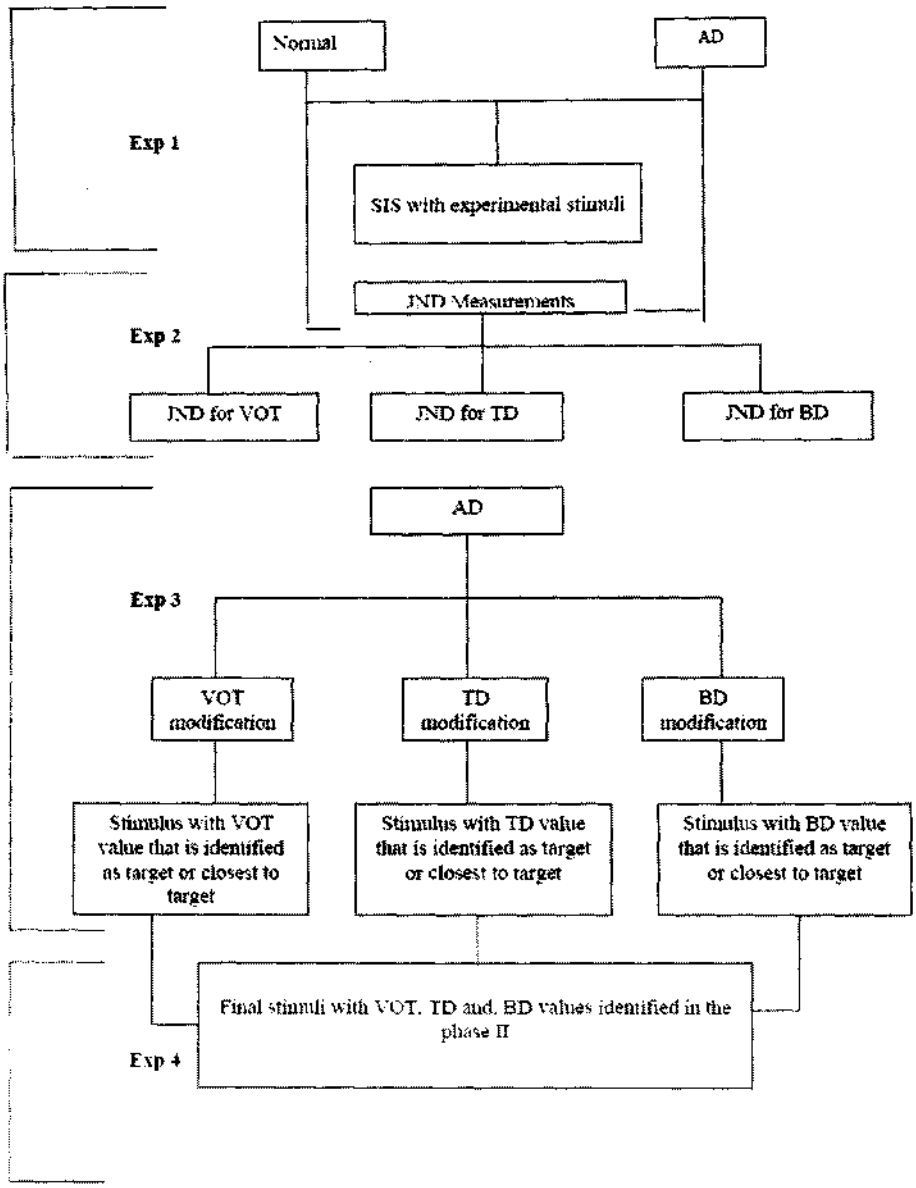


Figure 3.6 Block diagram of the sequence of experiments in Phase II of the study

3.0 Phase III

Modulation detection thresholds were measured by determining the sensitivity to sinusoidally amplitude modulated broadband noise as a function of modulation frequency. The main objective of this experiment was to measure the temporal resolution in individuals with auditory dys-synchrony and to see its relationship to speech perception in auditory dys-synchronics. Subjects in both auditory dys-synchrony and normal hearing groups participated in this phase.

3.1 Test stimuli

The stimuli consisted of unmodulated and sinusoidally amplitude modulated broadband noise of 500ms with a ramp of 2.5ms. The modulated signal was derived by multiplying the broadband noise by a dc-shiAed sine wave. The depth of modulation was controlled by varying the amplitude of modulating sine wave. Modulation depth for different stimuli varied between 0 to -30 dB (where 0 dB is equal to 100% modulation depth). Detection thresholds were measured for 5 modulated frequencies: 4 Hz, 16 Hz, 32 Hz, 64 Hz, 128 Hz, and 200 Hz. All the stimuli were generated using a 32 bit digital to analog converter at a sampling frequency of 44.1 kHz. Figure 3.7 shows one of the stimuli used in the study (16 Hz modulation with 100% and 50% modulation depth).

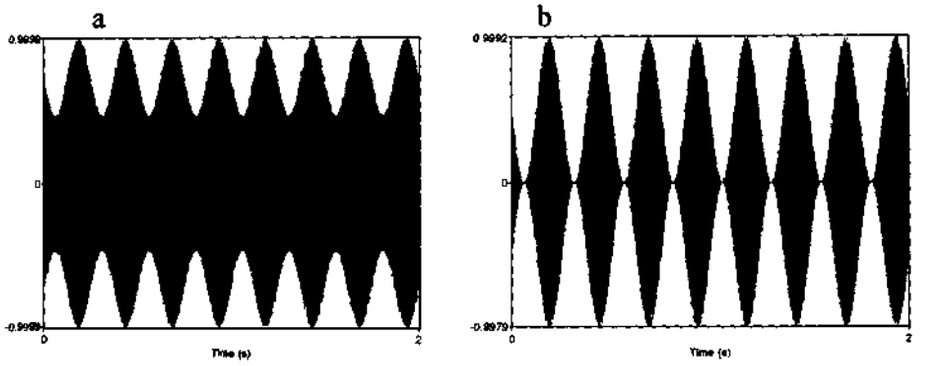


Figure 3.7. Sinusoidally amplitude modulated white noise at 16 Hz. (a) 50% modulation depth (b) 100% modulation depth.

3.2 Procedure

The procedure, test environment, instrumentation and presentation level was the same as in Experiment 2 of Phase II of the study. The worst threshold that can be obtained corresponds to a modulation depth of 0 dB (100% modulated noise). The presentation level was changed in all the subjects at least at one modulation frequency and modulation detection threshold was rechecked to ensure that subjects were not using loudness judgments.

3.3 Analyses

The following analysis were made:

a) Estimation of the prevalence of the problem of auditory dys-synchrony in Indian population

b) Comparison of speech identification scores of normals and subjects with auditory dys-synchrony.

c) Between - group comparison of JNDs for the temporal parameters of TD, BD and VOT - for each of the 8 CV syllables.

d) Within - group comparison of JNDs for TD, BD, and VOT for each of the 8 syllables

e) Speech identification scores with modified TD, BD and VOT, in isolation and in combination.

f) Between - group comparisons of modulation detection thresholds for sinusoidally amplitude modulated white noise at 4 Hz, 16 Hz, 32 Hz, 64 Hz, 128 Hz and 200 Hz.

g) Correlation between modulation detection thresholds and speech identification scores obtained in various experiments in individuals with auditory dys-synchrony.

CHAPTER 4

RESULTS

The objectives of this study were to (a) estimate the prevalence of auditory dys-synchrony in the South Indian city of Mysore, (b) evaluate perception of temporal parameters of speech in individuals with auditory dys-synchrony, and (c) measure modulation detection thresholds for sinusoidally amplitude modulated white noise in individuals with auditory dys-synchrony. The study was conducted in 3 phases. The data was analyzed on SPSS (version 10) and FIX- a shareware developed by the Department of Linguistics of the University College of London. Following analysis were carried out:

a) Prevalence of the problem of auditory dys-synchrony in Indian population.

b) Comparison of speech identification scores of normals and subjects with auditory dys-synchrony.

c) Between-group and within-group comparisons of just noticeable differences (JND) for the three temporal parameters of speech, namely, transition duration (TD), burst duration (BD) and voice onset time (VOT).

d) Speech identification scores in persons with auditory dys-synchrony with lengthened TD, BD and VOT, in isolation and in combination.

e) Between-group comparisons of modulation detection thresholds for sinusoidally amplitude modulated white noise at 4 Hz, 16 Hz, 32 Hz, 64 Hz, 128 Hz and 200 Hz.

f) Correlation between modulation detection thresholds and speech identification scores in individuals with auditory dys-synchrony.

1.0 Phase1

1.1 Prevalence

61 of 21,236 hearing impaired were identified to have auditory dys-synchrony following the criteria of Starr, Picton and Kim (2000) which means that 1 out of 348 hearing impaired had auditory dys-synchrony. However, the prevalence was 1 in 183 (0.53%) when only individuals with permanent sensori-neural hearing loss were considered (61 of 11,205).

The average age of onset was 16 years (range 1 to 31 years). 59% of the 61 patients with auditory dys-synchrony had onset of the problem when they were between the age of 14 to 24 years (Figure 4.1). 81% of this auditory dys-synchrony population reported their hearing problem as progressive. No specific etiology could be traced in most of these patients. Two patients had a neurodegenerative disorder and 2 patients reported that the problem started after they gave birth to their first child.

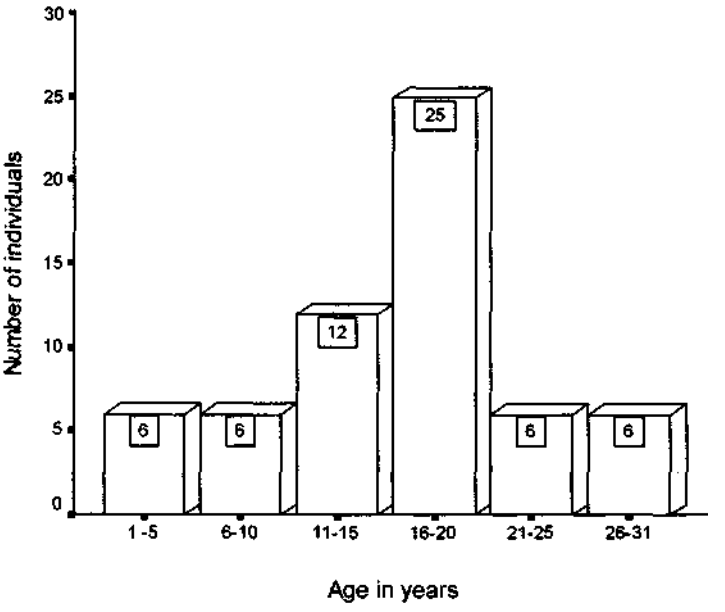


Figure 4.1. Age of onset of auditory dys-synchrony (N = 61).

Only two patients with auditory dys-synchrony demonstrated abnormal MRI while the remaining presented normal clinical neurological and imaging studies. All patients showed normal findings on otologic examination. There was no evidence of any middle ear disorder. 50% of the auditory dys-synchrony population reported bilateral tinnitus while 8 patients reported of vertigo. All patients who had vertigo complained of tinnitus also.

1.2 Puretone Audiometry

The auditory problem was bilateral in all the 61 patients though asymmetry was observed in some individuals. Females and males were affected in the ratio of 2:1. 54 of the 61 patients came from poor socioeconomic strata of the society (monthly income Rs < 2000 or < US$ 42).

Of the 61 patients, 26 showed peaked audiograms (sharp peak at a single frequency with worsening of thresholds at immediately adjacent frequencies), 11 showed flat audiograms (a raise or fall of <5 dB per octave), 11 showed a rising configuration (a decrease of 5 dB or more in thresholds per octave), 8 showed saucer-shaped audiograms (a loss of 20 dB or more at extreme frequencies compared to middle frequencies) and 3 showed sloping audiograms (an increase of 5-12 dB in thresholds per octave) in both ears (Figure 4.2). 77% of the patients who showed peaked audiogram, showed a peak at 2000 Hz. Puretone thresholds were not available in two children below 2 years of age.

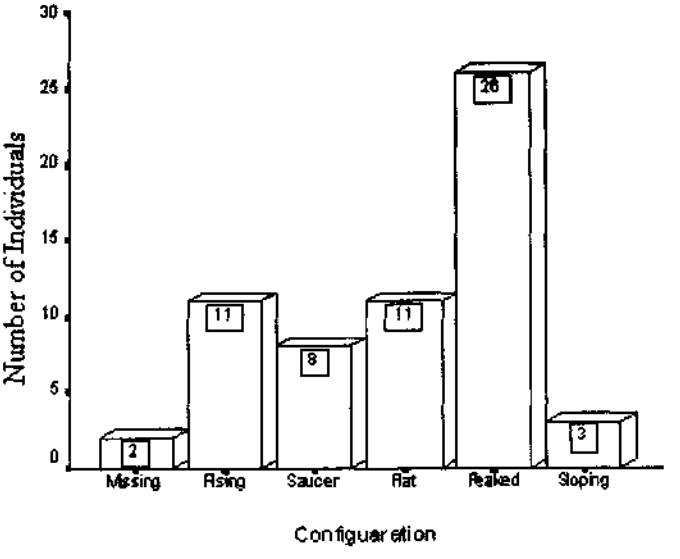


Figure 4.2. Audiogram configurations in individuals with auditory dys-synchrony.

Degree of hearing loss varied from mild to severe. This classification may not be valid in most of the cases because not only were the responses inconsistent, but also a majority of them had peaked audiograms. Responses of 40% of the patients were judged as inconsistent because thresholds varied by more than 10 dB within a test session. Paired two-tailed Y test showed no significant difference between the two ears with respect to 3-frequency pure tone average ($t = 0.5$, $p > 0.05$).

1.3 Speech Perception

Speech perception abilities of patients varied from no measurable speech identification score to 90% speech identification score. Figures 4.3a and 4.3b show the distribution pattern of speech identification scores for the right and the left ear, respectively. It is clear from Figures 4.3a and 4.3b that around 60% of the patients did not have measurable speech identification scores.

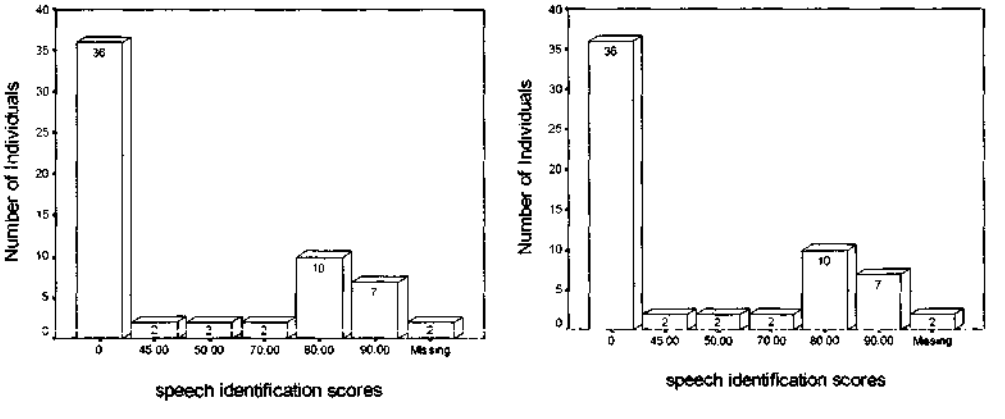


Figure 4.3. Speech identification scores in the right (4.3a) and the left (4.3b) ears.

Relationship between speech identification scores and the results of other auditory assessments was examined in all the patients. Patients' thresholds at each of the audiometric frequencies (octave frequencies between 250 Hz and 8 kHz) were compared with speech perception score by computing a Pearson's Product Moment Correlation between the two factors. There was a significant negative correlation between the two variables at all the frequencies except at 8 kHz. Table 4.1 shows the correlation coefficient 'r' between thresholds at different frequencies and speech identification scores. The observed correlation may have been because of subjects who had no measurable speech perception abilities in the presence of mild to moderate hearing loss. A scatter plot was drawn between the two variables, at all frequencies, to verify the validity of correlations (Figure 4.4).

Table 4.1. Correlation matrix between speech identification scores and hearing thresholds.

	Frequencies					
	250 Hz	500 Hz	1 kHz	2 kHz	4 kHz	8 kHz
Correlation Coefficient	-0.634*	-0.518*	-0.828*	-0.759*	-0.32*	-0.4

* $p < 0.05$

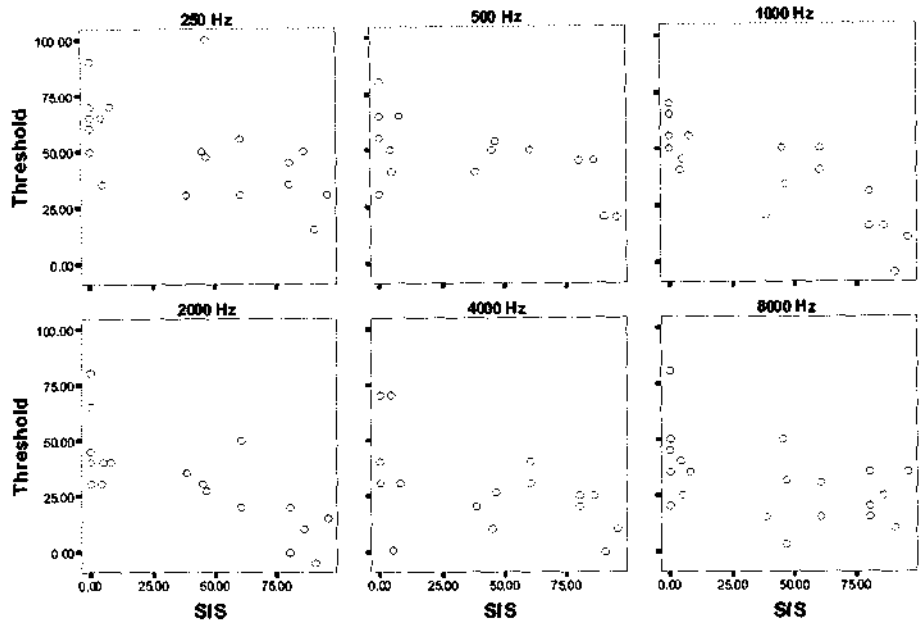


Figure 4.4. Relation between speech identification scores and thresholds at different audiometric frequencies.

An analysis of the scatter plots revealed a negative relation between hearing thresholds and the speech identification score only at 250 Hz and 500 Hz. Speech perception score was also compared with audiogram configuration. Figure 4.5 shows the mean speech identification score for each of the audiometric configurations described earlier. It is clear from this figure that patients who had peaked audiogram showed higher speech identification scores, in both ears, than patients with other audiometric configurations. But, any generalization of this result is hazardous because of the far too small number of patients, particularly those with rising and saucer shaped audiograms.

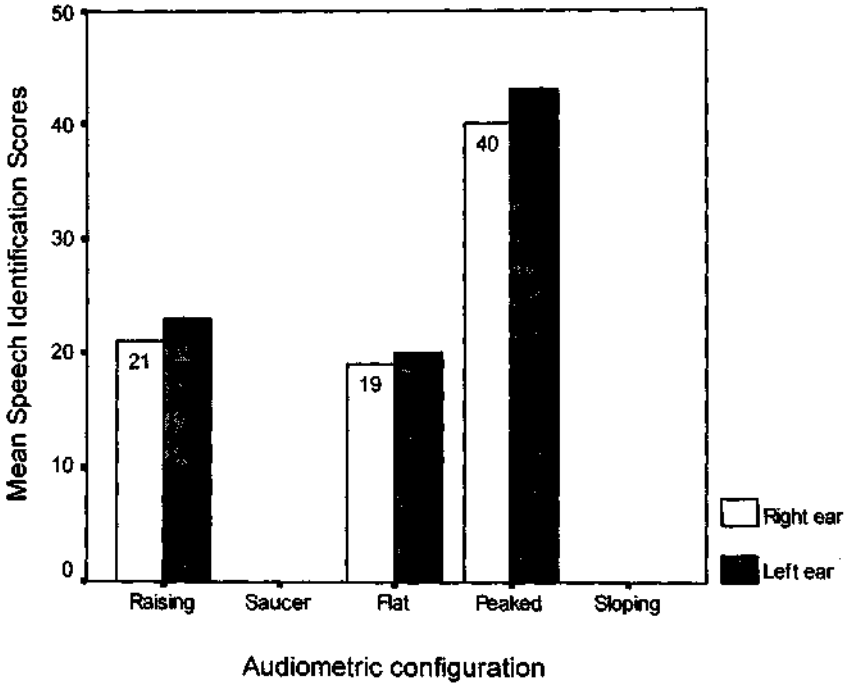


Figure 4.5. Speech identification scores in patients with different audiometric configurations.

1.4 Immittance Evaluation and ABRs

Immittance evaluation showed normal tympanograms. However, stapedial acoustic reflexes were absent in all patients. None of the patients showed ABRs for clicks in any ear.

1.5 Otoacoustic Emissions

Transient evoked otoacoustic emissions were bilaterally present in all the 61 subjects. The mean emission amplitude of TEOAE was 16.7 dB SPL

(SD = 3.9) for the right ear and 16.4 dB SPL (SD = 3.6) for the left ear. A two-tailed paired Y test failed to show any statistically significant difference between the right and the left ear in the amplitude of otoacoustic emissions (t = 1.23, p>.05). The relation between TEOAE amplitude and puretone thresholds at 500 Hz, 1 kHz, 2 kHz and 4 kHz was examined. Data from both the ears were combined as there was no statistically significant difference between the two ears for either TEOAE amplitude or puretone threshold. Pearson's Product Moment Correlation was not statistically significant between hearing threshold and TEOAE amplitude at any frequency. Correlation between the amplitude of TEOAEs and speech identification scores was not significant for either ear.

2.0 Phase II

Table 4.2 shows the audiological profile of 30 subjects with auditory dys-synchrony who participated in Phase II and Phase III of the study.

Table 4.2. Audiological profile of persons with auditory dys-synchrony.

SN	Age/sex	PTA (RE/LE)	SIS (RE/LE)	TEOAE	ABR	Acoustic Reflex	Efferent suppression (in dB)	Configuration
1	16/F	43/43	45/45	P	A	A	0.2	peaked
2	16/M	13/16	84/84	P	A	A	0.0	peaked
3	30/M	23/35	80/80	P	A	A	0.4	rising
4	24/F	31/46	38/36	P	A	A	0.0	rising
5	16/M	55/63	0/0	P	A	A	0.1	peaked
6	26/M	40/51	4/4	P	A	A	0.0	peaked
7	23/F	41/26	5/5	P	A	A	0.0	peaked
8	23/M	76/73	0/0	P	A	A	0.1	peaked
9	27/M	31/31	80/80	P	A	A	0.0	rising
10	23/M	23/23	86/84	P	A	A	0.3	peaked
11	24/M	53/53	8/8	P	A	A	0.0	peaked
12	25/F	43/43	68/68	P	A	A	0.0	peaked
13	28/M	31/31	80/80	P	A	A	0.3	rising
14	25/F	36/31	0/0	P	A	A	0.2	peaked
15	16/M	36/31	60/60	P	A	A	0.2	peaked
16	18/M	31/31	76/80	P	A	A	0.3	peaked
17	22/M	50/53	0/0	P	A	A	0.0	rising
18	24/F	3/3	90/90	P	A	A	0.0	peaked
19	23/M	15/13	95/95	P	A	A	0.2	peaked
20	26/M	56/65	0/0	P	A	A	0.1	peaked
21	20/F	48/51	0/0	P	A	A	0.2	rising
22	18/F	43/43	26/40	P	A	A	0.3	peaked
23	16/F	13/16	84/84	P	A	A	0.1	peaked
24	18/M	25/25	82/82	P	A	A	0.1	rising
25	19/M	31/46	38/36	P	A	A	0.0	rising
26	30/M	52/63	0/0	P	A	A	0.0	peaked
27	28/M	40/51	4/4	P	A	A	0.0	peaked
28	27/F	31/26	6/6	P	A	A	0.0	peaked
29	22/F	76/26	0/0	P	A	A	0.0	peaked
30	20/M	31/31	62/80	P	A	A	0.1	rising

RE - Right ear, LE - Left ear, SIS - Speech identification scores,
P - present, A - Absent.

2.1 Experiment 1: Speech Identification Scores with Unmodified Stimuli

All subjects in the normal hearing group got speech identification scores of 100% for the unmodified stimuli. The speech identification scores of subjects in the auditory dys-synchrony group ranged from 0% to 87% with a mean of 28% and an SD of 25.8%. Independent sample Y test showed a significant difference between the mean speech identification scores of the two groups ($t = 15.19$, $p < 0.01$).

2.2 Experiment 2: JNDs for Temporal Parameters of Speech

Repeated measures ANOVA revealed a highly significant difference between the JNDs of normals and auditory dys-synchronics for each of the three temporal parameters of speech, namely, transition duration [$F(1, 63) = 4471$, $p < 0.01$], burst duration [$F(1, 63) = 1389$, $p < 0.01$] and voice onset time [$F(1,63) = 2191$, $p < 0.0l$]. Independent sample Y tests were run for the significance of difference in mean JNDs for each of the CV syllables and separately for TD, BD and VOT. The questions asked were: (a) whether the JND for TD for /ba/ was significantly different between normals and auditory dys-synchronics, (b) whether the JND of VOT of /pa/ was significantly different between the two subject groups, and so on. Thus, a total of 24 Y tests were run. The results revealed a significant difference between the two groups in the JNDs for each of 8 CV syllables and for each of the temporal parameters. These multiple Y tests will not increase type I error because the

significant main effect at 0.01 level was established by repeated measure ANOVA for each of the temporal parameters of speech. Figures 4.6, 4.7 and 4.8 show mean JNDs for the 3 temporal parameters tested with 95% confidence interval error bars, between the two groups, across all the eight speech sounds.

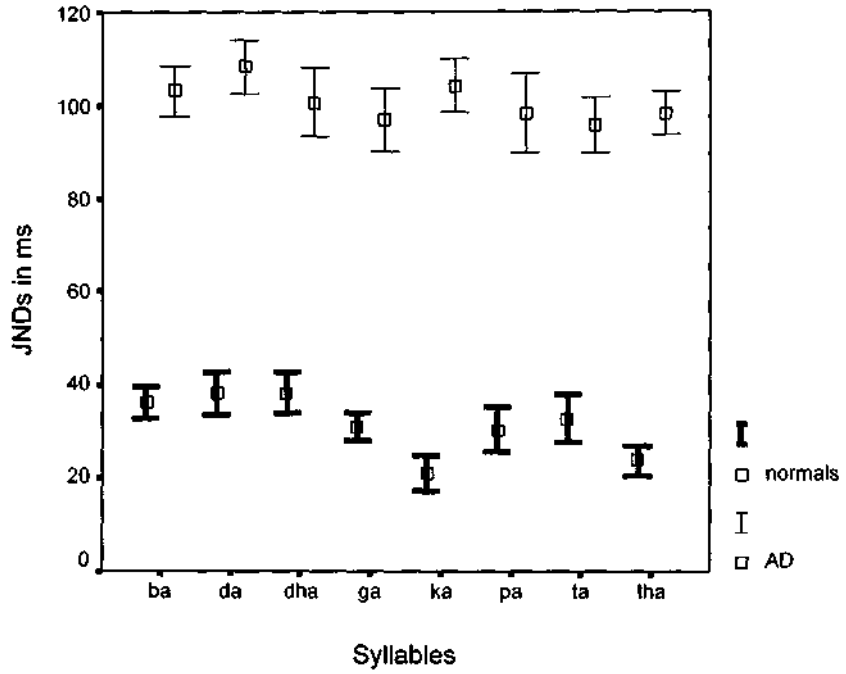


Figure 4.6. JND for transition duration for the two groups of subjects for each of the eight CV syllables. Error bars show 95% confidence interval. AD = auditory dys-synchrony.

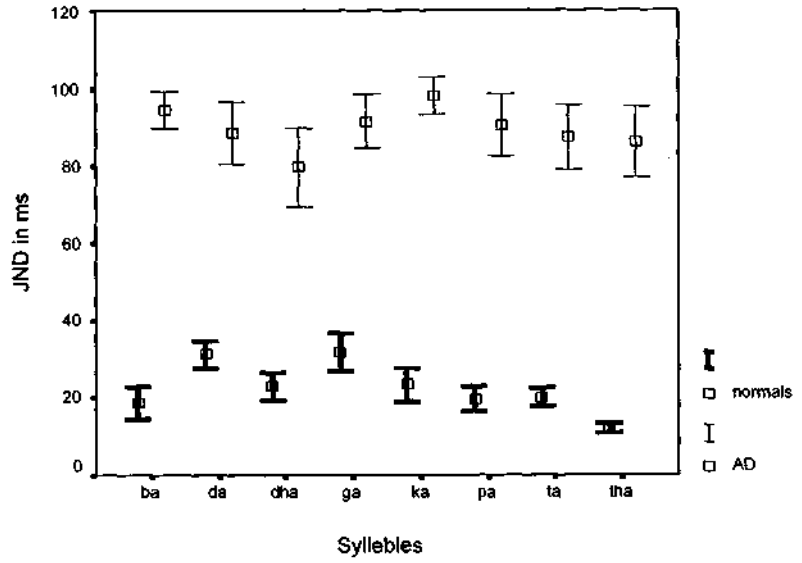


Figure 4.7. JND for burst duration for the two groups of subjects for each of the eight CV syllables. Error bars show 95% confidence interval. AD = auditory dys-synchrony.

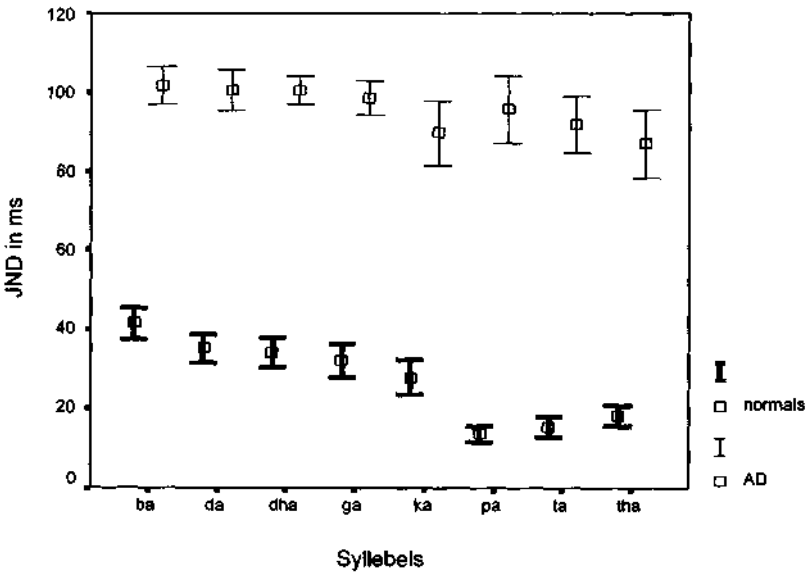


Figure 4.8. JND for voice onset time for the two groups of subjects for each of the eight CV syllables. Error bars show 95% confidence interval. AD - auditory dys-synchrony.

The difference in mean JND between the eight CV syllables, within each group, was analyzed with repeated measures ANOVA. The results of these analyses are described below:

2.2.1 Normal Hearing Group

Table 4.3 shows the mean and standard deviation of JNDs for the three temporal parameters of speech in normal hearing individuals. Repeated measures ANOVA revealed a significant effect of stimulus on the JNDs for BD [$F(7,203) = 18.5$, $p < 0.05$], TD [$F(7,203) = 13.4$, $p < 0.05$] and VOT [$F(7,203) = 44.4$, $p < 0.05$]. Bonferroni's pair wise comparison was carried out to analyze the interaction effects. Questions answered were: (a) whether the mean JND for TD for syllable /ba/ was significantly different from the JND for TD of other seven CV syllables; (b) whether the mean JND for TD of syllable /da/ was significantly different from the JND for TD of other seven CV syllables, and so on. Tables 4.4 a, b and c show results of this comparison.

Table 4.3. Mean JNDs, and standard deviation (SD) for TD, BD, and VOT in normal hearing group.

Stimulus	Parameter	Mean (in ms)	SD (in ms)	Range
/ba/	BD	18.6	11.3	10-40
	TD	36.1	8.7	20-50
	VOT	41.5	10.4	20-60
da/	BD	31.3	9.7	10-50
	TD	38.2	12.3	20-60
	VOT	35.1	9.9	20-60
/dha/	BD	22.8	9.4	10-50
	TD	32.8	11.4	20-60
	VOT	34.0	10.2	10-50
/ga/	BD	31.7	13.4	10-50
	TD	31.0	7.5	20-40
	VOT	32.1	11.1	10-50
/pa/	BD	19.6	8.0	10-40
	TD	30.3	12.4	10-50
	VOT	13.7	5.3	10-30
/ka/	BD	23.3	11.4	10-50
	TD	21.3	9.9	10-40
	VOT	27.9	11.8	10-50
/ta/	BD	20.0	6.4	10-30
	TD	32.6	13.7	10-60
	VOT	15.5	6.8	10-30
/tha/	BD	12.1	3.6	10-20
	TD	23.7	8.6	10-50
	VOT	18.2	6.7	10-30

Table 4.4a. Results of Bonferroni's pair wise comparison of JDNs for TD.

	ba	da	dha	ga	ka	pa	ta	tha
ba		NS	NS	NS	**	NS	NS	**
da			NS	NS	**	NS	NS	**
dha				NS	**	NS	NS	**
ga					**	NS	NS	NS
ka						**	NS	**
pa							NS	NS
ta								**
tha								

*p<0.01

NS = difference not statistically significant

Table 4.4b. Results of Bonferroni's pair wise comparison of JDNs for BD.

	ba	da	Dha	ga	ka	pa	ta	tha
ba		**	NS	**	NS	NS	NS	NS
da			NS	NS	**	**	**	**
dha				**	NS	NS	NS	**
ga					**	**	**	**
ka						NS	NS	**
pa							NS	**
ta								**
tha								

*p<0.01

NS = difference not statistically significant

Table 4.4c. Results of Bonferroni's pair wise comparison of JDNs for VOT.

	ba	da	dha	ga	ka	pa	ta	tha
ba		NS	**	**	**	**	**	**
da			NS	NS	NS	**	**	**
dha				NS	NS	**	**	**
ga					NS	**	**	**
ka						**	**	**
pa							NS	NS
ta								
tha								

*p<0.01

NS = difference not statistically significant

2.2.2 Auditory Dys-synchrony Group

Table 4.5 shows the mean and standard deviation of JNDs for the eight speech sounds and for each of the temporal parameters: TD, BD and VOT, in individuals with auditory dys-synchrony. A repeated measures ANOVA showed no significant effect of stimulus on JNDs of TD [F (7, 238) = 2.14, p> 0.05] and of VOT [F (7, 238) = 2.4, p> 0.05]. However, JNDs for BD significantly differed between the stimuli [F (7, 238) = 3.62, p>0.05]. Bonferroni's pairwise comparison showed that the mean JNDs for BD were

significantly different between the stimuli 'ba-dha' and 'dha-pa'. Figure 4.9 shows error bars of JNDs for burst duration, for the three syllables of /ba/, /dha/ and /pa/. It is clear from the figure that JNDs for these three stimuli overlap and that the observed statistical difference may be because of extreme scores.

Table 4.5. Mean JND and standard deviation (SD) for TD, BD and VOT of the subjects in auditory dys-synchrony group.

Stimulus	Parameter	Mean (in ms)	SD (in ms)	Range
/ba/	BD	95.5	14	40-100
	TD	103.1	16	60-150
	VOT	101.7	13.04	60-150
da/	BD	88.5	23.4	30-100
	TD	108.2	16.8	70-150
	VOT	100.5	14.7	60-150
/dha/	BD	82	26.8	20-100
	TD	100.5	21.4	40-150
	VOT	100.5	9.9	70-150
/ga/	BD	91.4	20.3	30-100
	TD	96.8	19.8	50-150
	VOT	98.5	13	60-150
/pa/	BD	90.5	23.5	10-100
	TD	98.2	24.6	40-150
	VOT	95.7	24.7	40-150
/ka/	BD	98.2	23.5	10-100
	TD	104.2	24.6	40-150
	VOT	89.7	24.7	40-150
/ta/	BD	87.2	24.1	30-100
	TD	95.4	17.5	50-150
	VOT	91.7	20.9	30-100
/tha/	BD	86.2	26.4	30-100
	TD	98.2	13.1	50-130
	VOT	86.8	24.8	30-100

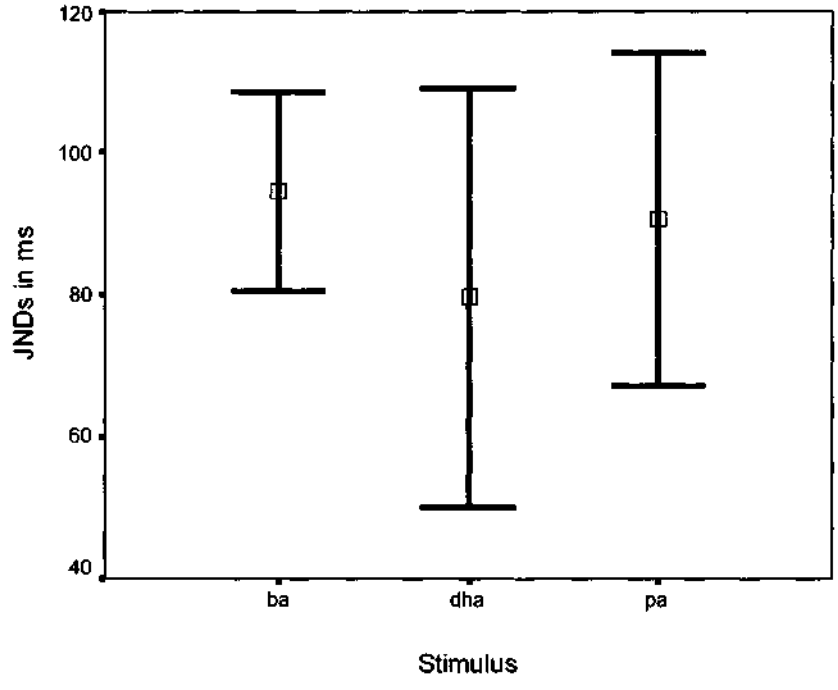


Figure 4.9. JNDs for burst durations for /ba/, /da/ and /dha/. Error bars show mean + 1SD.

2.3 Experiment III: Speech Identification Scores for Temporally Modified Stimuli (in isolation)

2.3.1 Speech Identification Scores

Each CV syllable was presented 10 times in each testing condition. This resulted in a total of 320 presentations per subject (8 CV syllebels X 4 conditions X 10 presentations). The four testing conditions were unmodified, transition duration modified, burst duration modified and voice onset time modified. Figures 4.10 (a-d) show the speech identification scores obtained by subjects in different testing conditions. Speech identification scores varied

between 0 to 87% for unmodified stimulus, 0 to 100% for stimuli with modified transition duration and VOT, and between 0 to 50% for stimuli with modified burst duration. Other important observations from Figures 4.10 (a-d) are as follows:

a) 11 patients with auditory dys-synchrony had 0% speech identification score for unmodified stimuli. Only four patients had speech identification scores more than 50%.

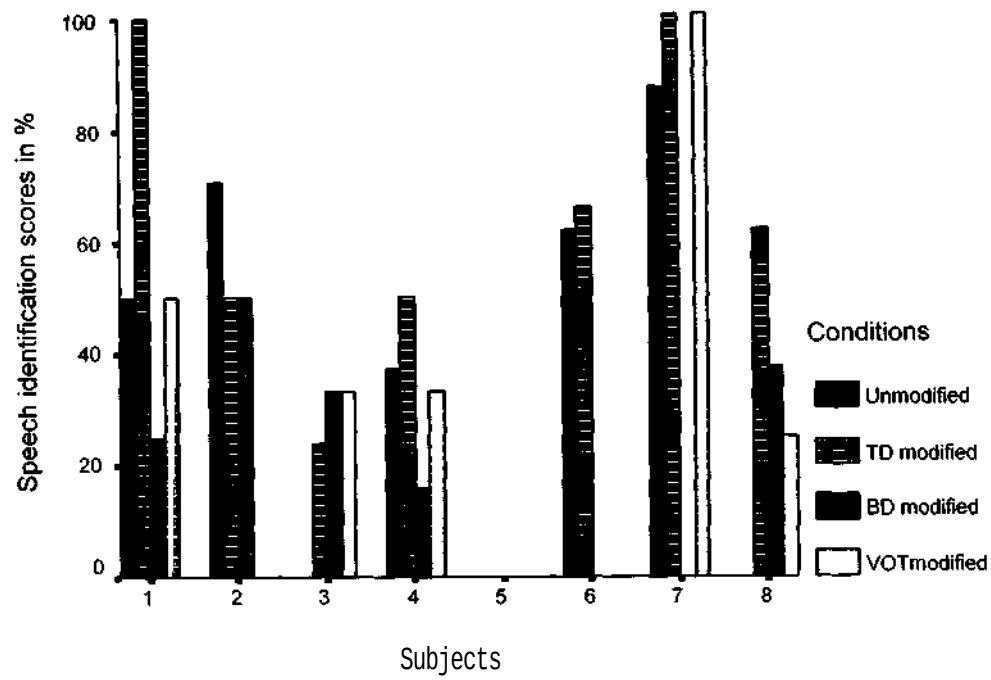
b) Five subjects with auditory dys-synchrony had 100% speech identification scores for stimuli with modified TD

c) Modification of burst duration resulted in less than 50% speech identification scores in all the subjects. When a stimulus was modified for TD, it means that other temporal parameters like BD and VOT was as in original unmodified stimuli.

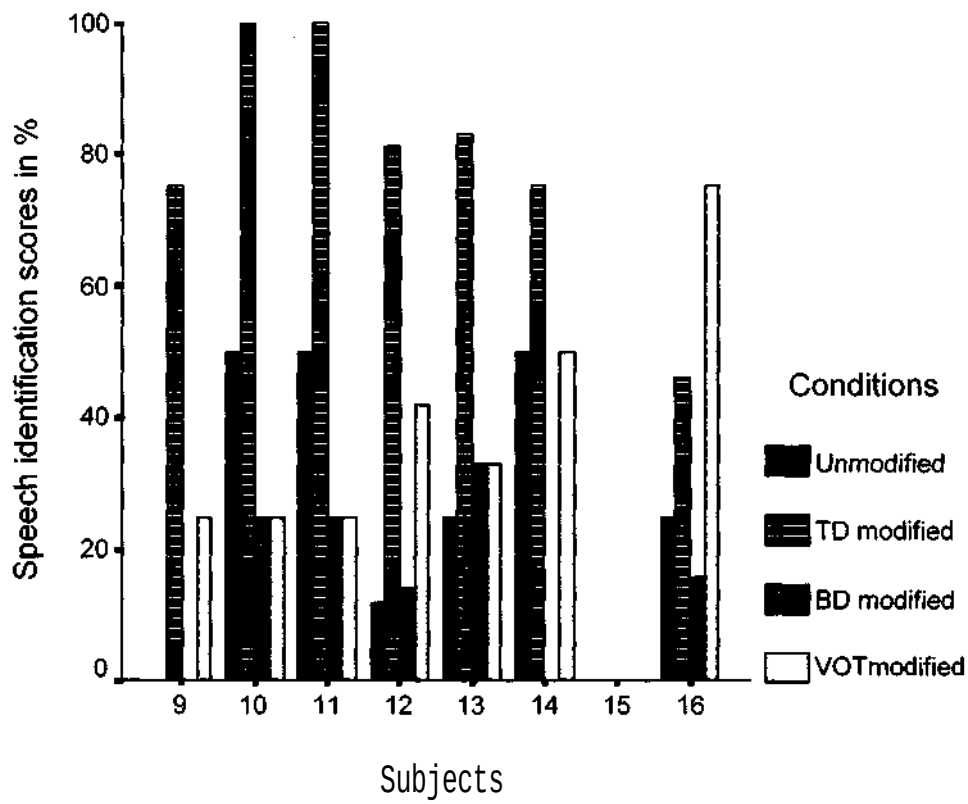
d) Modification of voice onset time resulted in 100% speech identification scores in one subject while six subjects obtained more than 50% speech identification scores.

Figure 4.11 shows the mean speech identification scores in the four testing conditions. Modification of transition duration resulted in better perception of CV syllables followed by modification of voice onset time and burst duration.

a


Speech identification scores in %
100
80
60
40
20
0
1
2
3
4
5
6
7
8
Subjects
Conditions
Unmodified
TD modified
BD modified
VOTmodified

b


Speech identification scores in %
100
80
60
40
20
0
9
10
11
12
13
14
15
16
Subjects
Conditions
Unmodified
TD modified
BD modified
VOTmodified

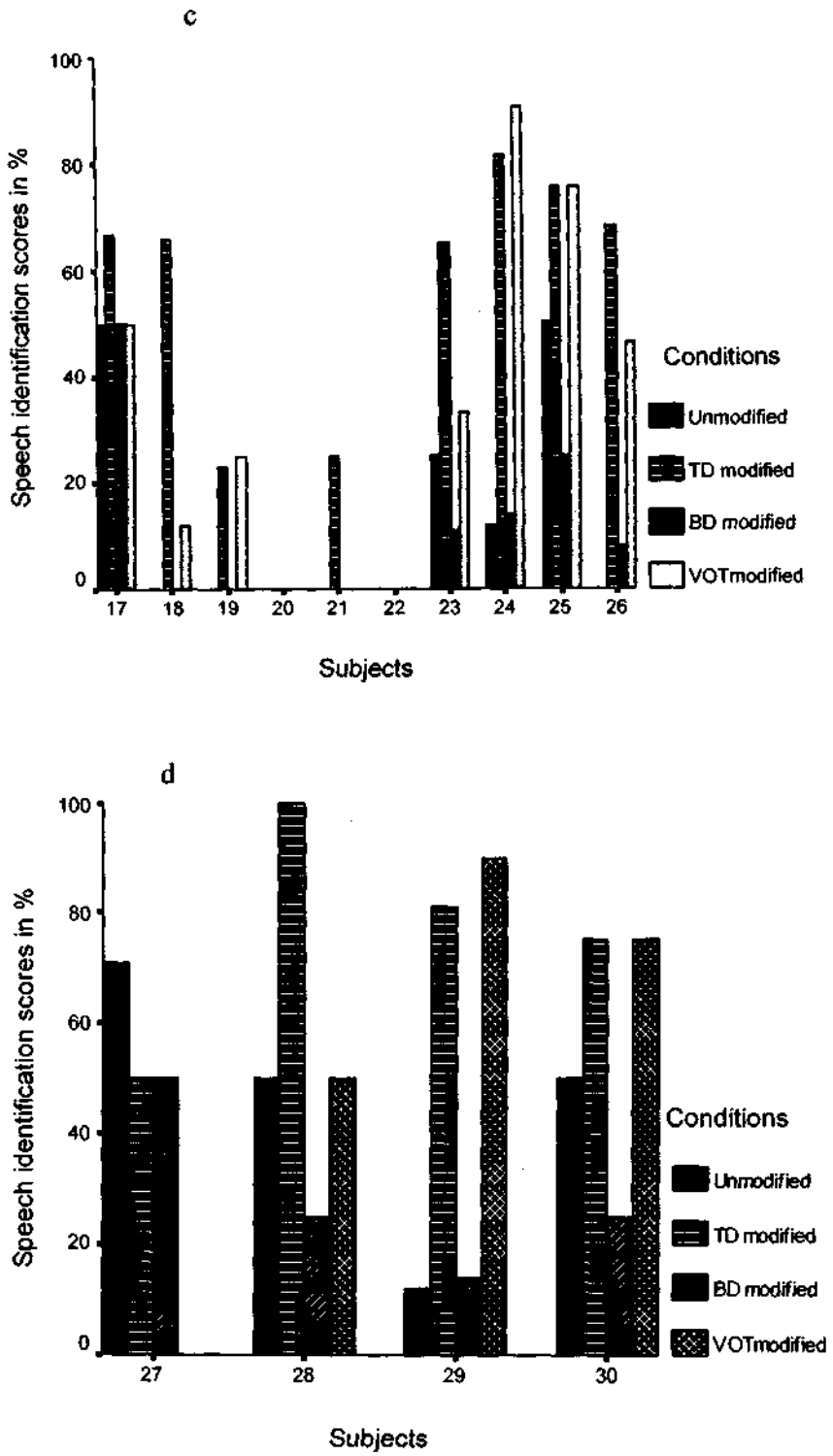


Figure 4.10 (a-d). Speech identification scores of individual subjects under different conditions.

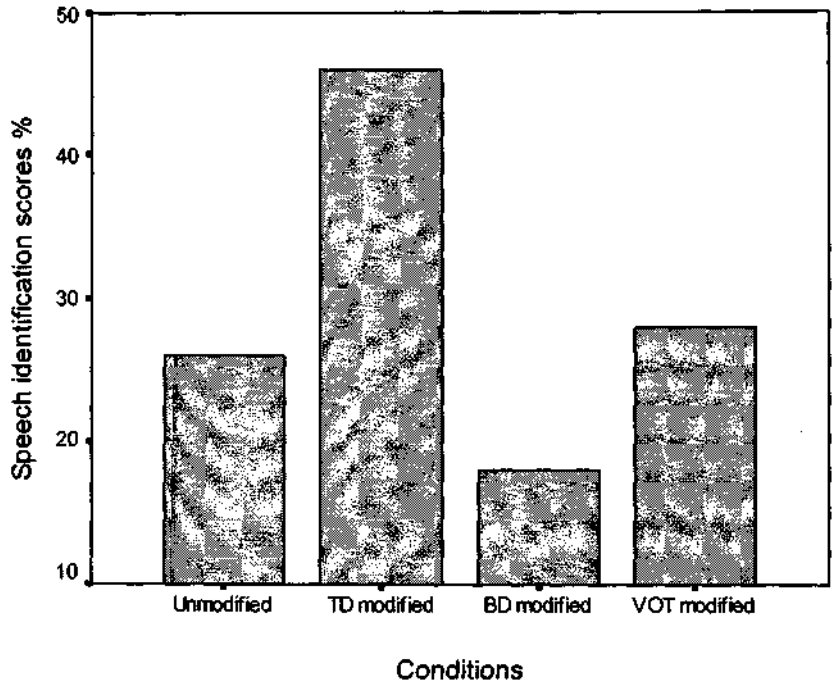


Figure 4.11. Mean speech identification scores under different testing conditions.

2.3.2 Speech Identification Scores for Each Speech

Figure 4.12 shows mean speech identification scores for the unmodified and the temporally modified CV syllables (modified for TD, BD and VOT). As can be seen modification of transition duration resulted in better speech identification scores for all syllables except for /ga/. Modification of voice onset time resulted in better speech identification scores for stimulus /ga/. Modification of transition duration resulted in speech identification scores of greater than 50% for all speech sounds other than /ka/, /pa/and/tha/.

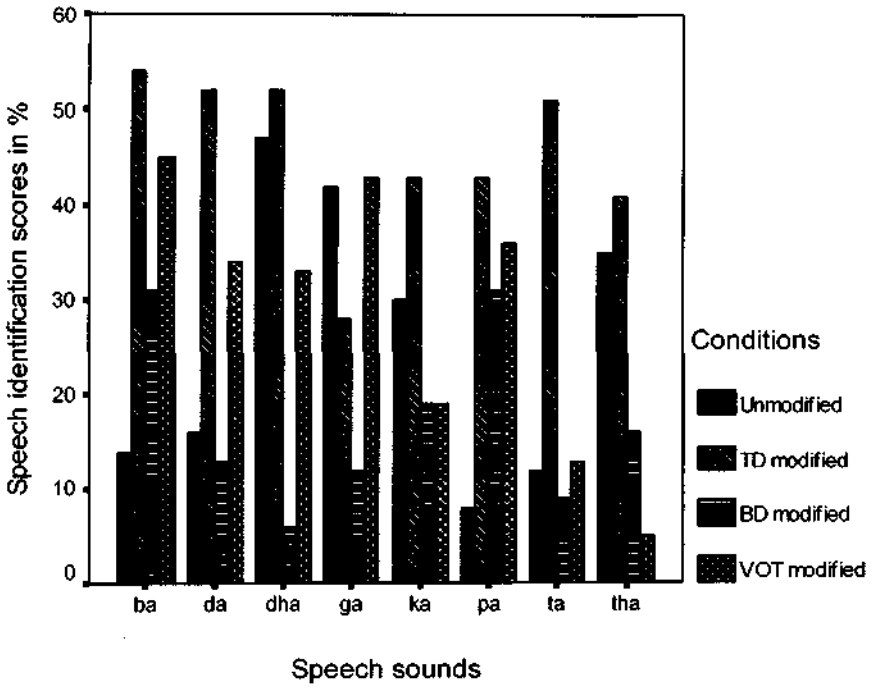


Figure 4.12. Speech identification scores for modified stimuli.

2.3.3 Stimulus- Response Matrices

Separate stimulus-response matrices were constructed for unmodified as well as for each of the three modified conditions. In these matrices (Tables 4.6-4.9), the number in each cell is the frequency with which each stimulus-response pair occurred. The number of correct responses may be obtained by totaling the frequency along the main diagonal (by adding the bold numbers). In each of the conditions, the stimulus concerned was presented 10 times. This resulted in 300 presentations of each stimulus, in the unmodified condition, for 30 subjects. However, in all the three modified conditions, the

total number of stimulus presentations was always less than 300. This is because number of speech stimuli that the subjects could identify in the unmodified condition was not considered in future experiments. Therefore, the total number of presentations was different for different stimuli in the modified conditions, but they remained constant across modified conditions. Hence, any number in the stimulus-response matrix can be directly compared between each of the three modified conditions, but none of these should be directly compared to any in the unmodified condition. Moreover, the row total for each stimulus in these matrices is less than the total number of presentations because some subjects who could not label the stimuli (identify) correctly, reported it as noise. These responses were not considered while constructing the stimulus-response matrix.

2.3.4 Stimulus- Response Matrix for Unmodified Speech Sounds

Table 4.6 is the stimulus-response matrix for the unmodified condition. A preliminary look at the matrix reveals that speech identification scores did not exceed 50%. Other observations from this matrix are as follows:

a) Individuals with auditory dys-synchrony perceived /dha/ better in relation to other speech sounds. This was closely followed by the identification of syllable /ga/.

b) No consistent grouping among phoneme categories was evident in the stimulus-response matrix. The exceptions were /ba/ which was

frequently confused with its unvoiced cognate /pa/, and /dha/ which was confused with /ga/.

c) The phonemes /ta/ and /pa/ were rarely identified correctly.

Table 4.6. Stimulus response matrix for unmodified speech sounds. Bold numbers indicate correct responses. The number in each cell reflects the combined responses from all the subjects.

		Response								
		ba	da	dha	ga	ka	pa	ta	tha	No response
Stimulus	ba	42		12	12	12	96	60	24	42
	da		48	60	24	12	12	36		108
	dha	12		**143**	63					82
	ga	12	60	24	**128**	15	12			49
	ka		48	12	24	**100**	12	24	24	56
	pa				17	24	24	48	72	115
	ta	12	48	12	36		12	36	50	94
	tha		31			24	42	12	**105**	86

2.3.5 Stimulus Response Matrix for Speech Sounds Modified for Transition Duration

Stimulus response matrix for stimuli modified for transition duration is shown in Table 4.7

Table 4.7. Stimulus response matrix for speech sounds modified for transition duration. Bold numbers indicate correct responses. The number in each cell reflects the combined responses from all the subjects.

		Response									
		ba	da	dha	Ga	ka	pa	Ta	tha	No responses	Total number of presentations
stimulus	ba	**144**	12	12	5		36	12	11	32	264
	da		**139**	52	12				12	49	264
	dha			**95**					3	82	180
	ga		36	24	**55**	18				59	192
	ka		12			96	24	14		74	220
	pa				12		**115**	24	24	89	264
	ta		24	24			12	**141**	13	62	276
	tha					4		12	90	110	216

The matrix reveals that lengthening transition duration resulted in greater number of correct responses when compared to unmodifiedstimuli. This was true for all the speech sounds. This resulted in correct identification of more than 50% in respect of syllables /ba/, /da/, /dha/ and /ta/. It is also observed that:

a) confusion of /ba/ with /pa/, persisted, but with greatly reduced frequency, and

b) /dha/ continued to be confused with /da/,

2.3.6 Stimulus Response Matrix for Speech Sounds Modified for Burst Duration

Table 4.8 shows the stimulus-response matrix for stimuli modified for burst duration. Lengthening of burst duration improved identification of only /ba/ and /pa/ syllables compared to unmodified stimuli. However, frequency of correct identification did not cross 50% for even these two syllables. Lengthening of burst duration also resulted in

a) reduced number of responses. Many a time the subjects could not even label the speech sounds, and

b) increased confusion between speech sounds

Table 4.8. Stimulus-response matrix for speech sounds modified for the burst duration. Bold numbers indicate correct responses. The number in each cell reflects the combined responses from all the subjects.

		Response									
		ba	da	dha	ga	ka	pa	ta	tha	No responses	Total number of presentations
stimulus	ba	84		48	12		72	12		36	264
	da	48	36	12		12		24	24	108	264
	dha	36		12	24		12		12	84	180
	ga		36	12	24	24		12	12	72	192
	ka	12	12		12	42		24		118	220
	pa		12		12	12	84	24	48	72	264
	ta		12	36	12		12	24	12	168	276
	tha	12	12	12		12	12		36	120	216

2.3.7 Stimulus-Response Matrix for Speech Sounds Modified for Voice Onset Time

Table 4.9 shows stimulus-response matrix for stimuli whose VOT was lengthened. Lengthening voice onset time resulted in a decrease of 'across category' (voicing) confusions in comparison to 'within category' confusions. For instance, /ka/ was confused with other unvoiced sounds such as /pa/, /ta/, and /tha/, but never with any voiced sound. This is shown in dark and lightly shaded area in Table 4.9. It was also observed that

a) there were quite a few instances of voiced sounds being identified as unvoiced sounds, and

b) correct identification of /ta/, /tha/ and /ka/ was poorer compared to identification of other CV syllables.

Table 4.9. Stimulus-response matrix for speech sounds modified for voice onset time. Bold numbers indicate correct responses. The number in each cell reflects the combined responses from all the subjects.

	Response										
		ba	da	dha	ga	Ka	pa	ta	tha	No responses	Total number of presentations
stimulus	ba	12©	**12**	**12.**		24		48		48	264
	da	-	**90**	**36**				12	6	120	264
	dha	**12**		**60**						108	180
	ga		12	22	**84**		12			62	192
	ka					**42**	15	12	12	139	220
	pa					24	96	12	24	108	264
	ta			24	12		24	36	24	156	276
	tha			12	12	12	24	48	12	96	216

2.3.8 Sequential Information Transfer Analysis

A sequential information transfer analysis (Wang & Bilger, 1973) was performed on group data for each experimental condition to assess the amount

of information transfer from stimulus to response for a set of phonetic feature. This was done with 'Feature Information Xfer (FIX)' software from the Department of Linguistics, University College of London. Sequential Information Transfer Analysis (SINFA) is a method for determining the degree of information transfer from stimulus to response that is attributable to a particular feature. First, each feature's information transfer from stimulus to response is computed. Then a sequence of iterations is done in which one feature selected according to some criteria is partialled out per iteration by holding it constant. Typically, the feature that is transmitted to the maximum extent is held constant in subsequent iterations. Table 4.10 shows the different features that were assigned to speech sounds. The speech sounds have been classified based on the features of voicing and place of articulation. Thus the speech sounds were either voiced or unvoiced (voicing), and bilabial, or dental, or velar or alveolar (place).

Table 4.10. Features assigned to the different CV syllables.

	ba	da	dha	ga	ka	pa	ta	tha
Voicing	+	+	+	+	-			
Place	b	a d	v	v	b	a	d	

b = bilabial, a = alveolar, d = dental, v = velar

Table 4.11 shows the relative information transmitted in bits per stimulus for each feature. The maximum information that could be transmitted for the eight stimuli was three bits.

Table 4.11. Relative information transmitted for each feature

	Voicing (in bits)	Place (in bits)	Total transferred information (in bits)
Unmodified	0.129	0.186	0.812
TD modified	0.464	0.514	1.535
BD modified	0.133	0.317	0.731
VOT modified	0.416	0.346	1.322

The numbers in the 2nd and the 3nd columns represent the extent of information transmitted. 0 indicates no transmission of that particular feature and a value of '1' indicates maximum transmission of information. The numbers in the fourth column represent the total information transmitted. As said earlier information transmitted in this particular experimental paradigm is between 0 and 3 bits.

In general, the pattern of information transfer as shown by SINFA analysis was similar to the pattern found in the confusion matrices (Table 4.7 - 4.10). It was found through SINFA analysis that the stimuli modified for transition duration facilitated the highest information transfer followed by stimuli modified for VOT and burst duration, in that order. Some salient findings of SINFA analysis were as follows:

1. The modification of transition duration resulted in the transmission of place information to a greater extent than voicing information. This was also true with regard to stimuli modified for burst duration.

2. The chosen set of features could not account for the total information transferred from stimulus to response.

2.4 Experiment 4: Speech Identification with Temporally Modified Stimuli (in combination)

2.4.1 Speech Identification Scores

In this experiment speech identification testing was done with stimuli which had been modified to reflect the JNDs for TD, BD and VOT that had elicited the best speech identification scores in Experiment 3. These stimuli are labeled here as 'combined stimuli' for want of better nomenclature. Figure 4.13 (a & b) show individual speech identification scores for all the subjects for the combined stimuli. The last bar in Figure 4.13b shows the mean correct response of the auditory dys-synchrony group. The scores ranged from 0% to 100%. In none of the subjects*, combined modification reduced speech identification scores. Other observations from the data are as follows:

a) In 16 subjects, speech identification scores remained the same as those obtained for stimuli modified for TD. Speech identification scores of these subjects were absolutely same as those obtained with the stimuli modified for TD

b) In nine subjects, speech identification scores for combined stimuli were better than those resulted from modification of transition duration alone.

' In 12 subjects combined modification did not involve burst duration modification as it did not alter the speech identification scores compared to unmodified condition

c) 7 subjects had 100% speech identification scores on combined stimuli.

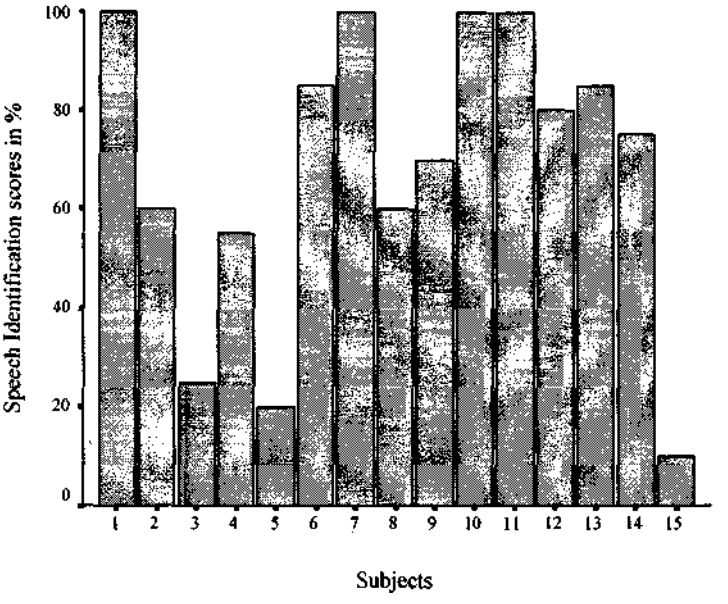


Figure 4.13a. Speech identification scores of individual subjects (1-8) for combined modification.

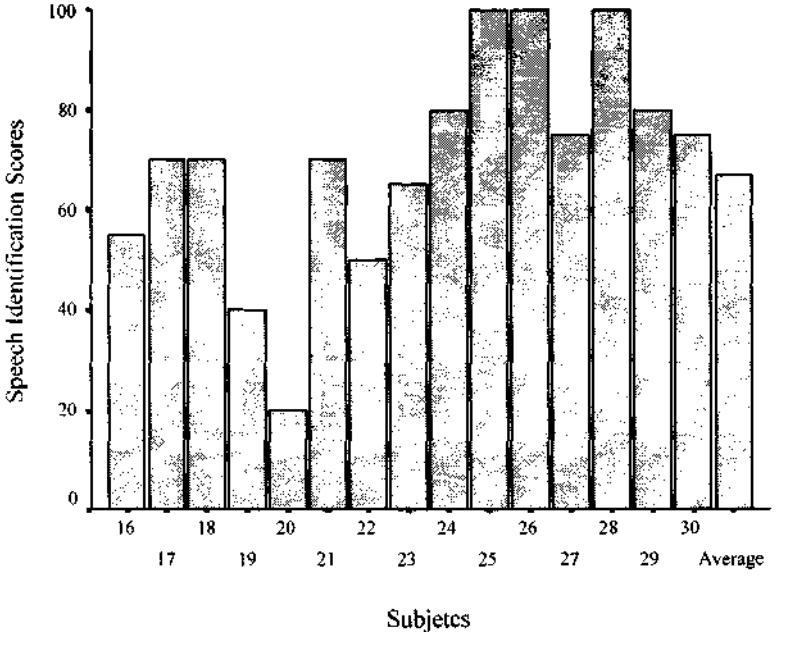


Figure 4.13b. Speech identification scores of individual subjects (16-30) combined modification.

The last bar in Figure 4.13b is the mean speech identification score for combined stimuli. A repeated measure ANOVA was performed for significance of difference of mean speech identification scores across the 5 conditions of speech (unmodified, TD modified, BD modified, VOT modified and optimally modified). Results showed that the main effect of stimulus modification on speech identification scores was statistically significant [F = (4, 132), 33.89, $p < 0.05$]. Bonferroni's post hoc analysis showed that modification of transition duration and 'combined modification' of all the three parameters resulted in significantly better speech identification scores than other modifications. Combined stimuli yielded better speech identification scores in many subjects, but the group effect was not statistically significant.

2.4.2 Speech Identification Scores for Each Speech Sound

Figure 4.14 shows identification scores for each speech sound for the combined modification in comparison with other conditions. The results observed were similar to those obtained with stimuli whose transition duration was modified, /ba/, /dha/ and /ta/ was identified more than 50% of the time. Identification scores for /ga/ was less than other speech sounds.

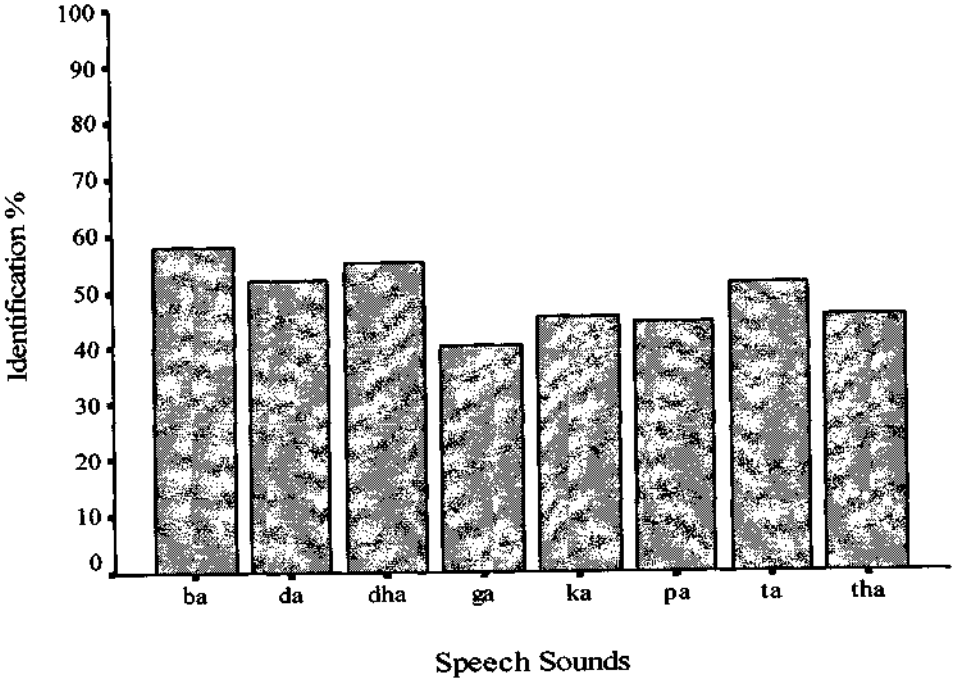


Figure 4.14. Identification scores for 8 CV syllables for optimally modified stimuli.

2.4.3 Stimulus-Response Matrix for Optimally Modified Speech Sounds

Table 4.12 shows the stimulus-response matrix for the optimally modified stimuli. It can be seen from Table 4.12 that the combined modification definitely increased the total number of correct responses. However, confusions between speech sounds remained at the same level as those for stimuli modified for TD (Table 4.7).

Table 4.12. Stimulus-response matrix for the optimally modified speech sounds. Bold numbers indicate correct responses. The number in each cell reflects the combined responses from all the subjects.

		Response									
		ba	da	dha	ga	ka	pa	ta	tha	No response	Total
Stimulus	ba	**152**	12		5	12	36	12	11	24	264
	da		137	48	12				12	55	264
	dha			98					2	80	180
	ga		24	24	72	30				42	192
	ka		12			96	24	12		76	220
	pa				12		115	24	24	89	264
	ta	12	12	23			12	**141**	12	64	276
	tha					3		12	97	168	216

Sequential information transfer analysis was carried on the stimulus response matrix shown in Table 4.12. Results of SINFA are shown in the Table 4.13.

Table 4.13. Information transmitted for each speech conditions. 1 to 4 are from Table 4.11 for comparison here.

		Voicing (in bits)	Place (in bits)	Total transferred information (in bits)
1	Unmodified	0.129	0.186	0.812
2	TD modified	0.464	0.514	1.535
3	BD modified	0.133	0.317	0.731
4	VOT modified	0.416	0.346	1.322
5	Combined modified	0.515	0.422	1.312

It can be seen from the table 4.13 that combined modification of speech stimuli (TD, BD and VOT) did not lead to the expected results speech identification. Information transfer for combined modification was at the same level as that for stimuli modified for VOT, and less than that for stimuli modified for TD alone.

3.0 Phase HI

3.1 Modulation Detection Thresholds

Figure 4.15 shows the temporal modulation transfer function (TMTF) for subjects with normal hearing and auditory dys-synchrony. Normal hearing listeners were most sensitive to slow temporal fluctuations but became less

and less sensitive for increase in fluctuation rates. A similar trend was noticed in individuals with auditory dys-synchrony. Average peak sensitivity of normal hearing listeners was -17.36 dB (SD: 1.2 dB). In contrast, average peak sensitivity for auditory neuropathy group was -6.2 dB (SD: 1.8 dB). At higher modulation frequencies, many of the subjects with auditory dys-synchrony (20 subjects) did not even detect modulation depth of 0 dB (100%). Figure 4.15 shows TMTF for subjects with normal hearing and auditory dys-synchrony.

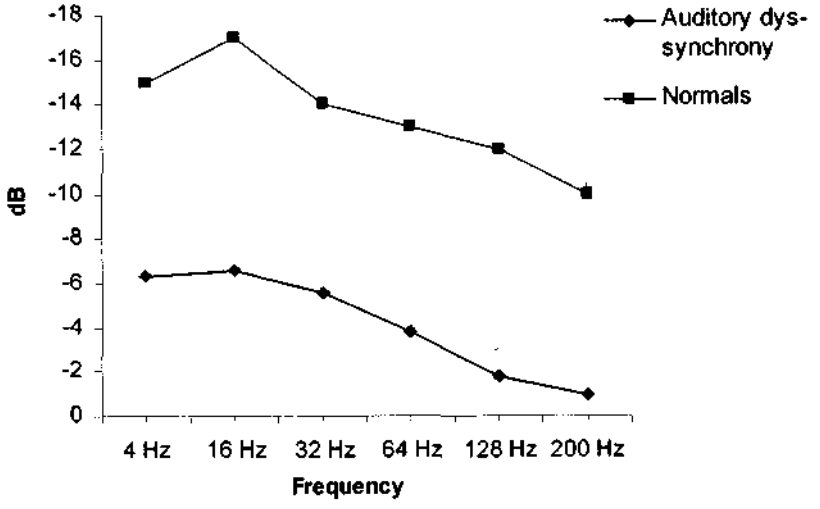


Figure 4.15 TMTF in normal hearing individuals and individuals with auditory dys-synchrony.

Pearson's Product Moment Correlation was computed between peak modulation detection thresholds and speech identification scores obtained for each acoustic modification. Table 4.14 shows correlation coefficient 'r' between peak modulation detection threshold and speech identification scores obtained in different conditions.

Table 4.14 Correlation matrix between peak modulation detection thresholds and speech identification scores obtained in different conditions.

Unmodified	TD lengthened	BD lengthened	VOT lengthened	combined modification
-0.566*	-.611*	-.272	-.241	-.3*

*p<0.05

There was a significant correlation between peak modulation detection threshold and speech identification scores obtained for unmodified, transition duration lengthened and combined modification stimuli. Figure 4.16 shows scatter plot and regression line between modulation detection threshold and speech identification scores obtained for different experimental conditions. It is clear from Table 4.14 and Figure 4.16 that speech perception abilities of individuals with auditory dys-synchrony were related to temporal resolution. Individuals with better modulation detection thresholds had better speech identification scores and benefited more from lengthening of transition than other subjects.

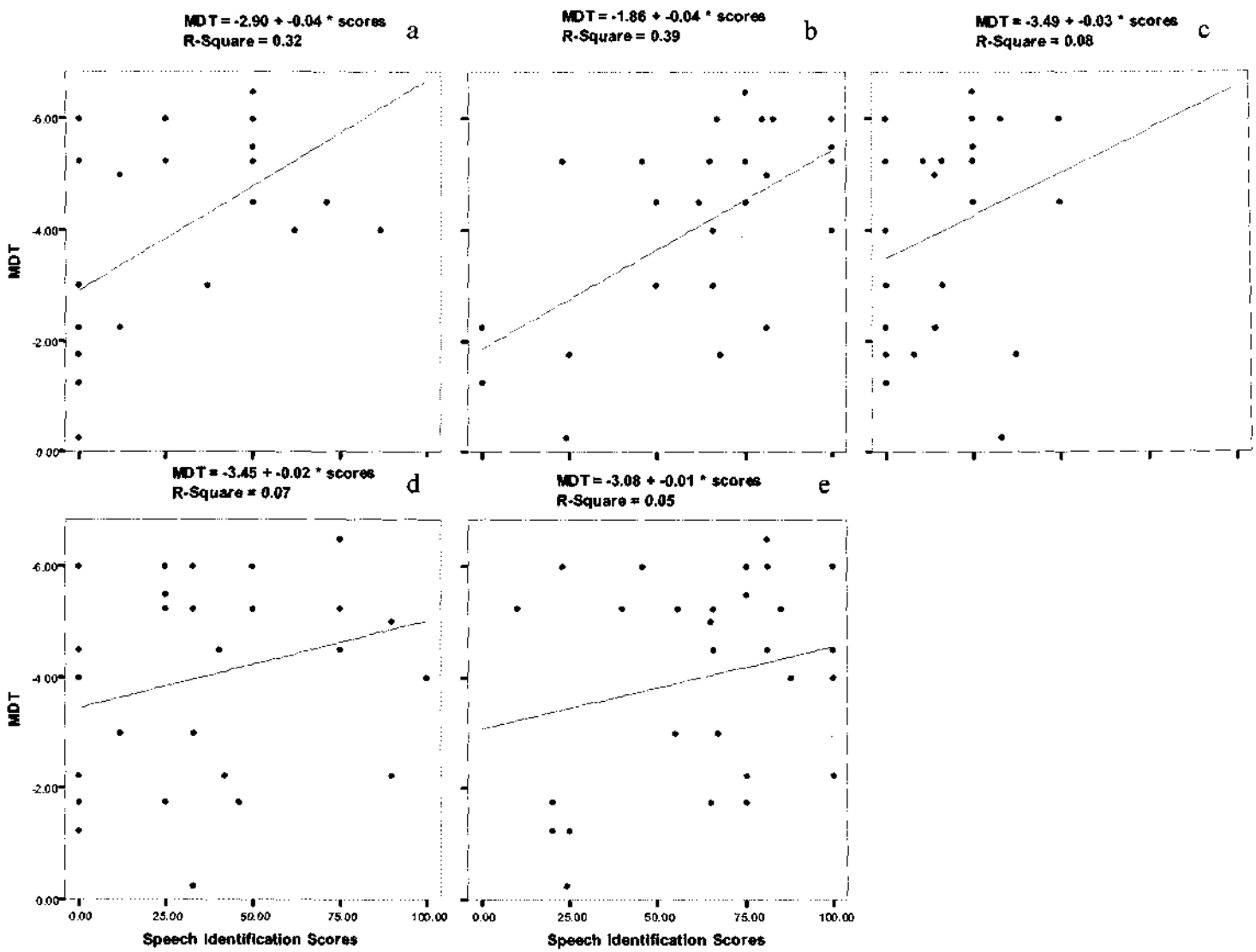


Figure 4.16. Scatter plot between peak modilation detection thresholds and speech identification scores obtained in unmodified (a), transition duarion modified (b), burst duration modified (c), VOT modified (d) and combined modification (e) conditions.

In summary, some important findings of this study are:

a) prevalence of auditory dys-synchrony among adults in India is about 0.53% of sensory neural hearing impaired and 0.28% when all types of hearing impairments are considered,

b) individuals with auditory dys-synchrony have severely affected temporal processing abilities as evident in elevated JND values for TD, BD and VOT,

c) no consistent error patterns between phoneme categories could be observed for unmodified speech sounds,

d) modification of transition duration resulted in significantly better speech perception than modification of either BD or VOT,

e) lengthening of VOT reduced across-category confusions (voiced vs unvoiced) between voiced and unvoiced speech sounds,

f) lengthening of BD did not result in improvement in speech perception and,

g) combined modification of TD, BD and VOT did not significantly improve speech identification compared to that for TD modification alone,

h) individuals with auditory dys-synchrony showed significantly poorer modulation detection thresholds than normal hearing subjects,

i) speech identification scores of individuals with auditory dys-synchrony was a significantly correlated with their modulation detection thresholds, and hearing sensitivity in the low frequency hearing region.

CHAPTER 5

DISCUSSION

1.0 Prevalence and Audiological Profile of Auditory Dys-synchrony

In the first phase, prevalence of auditory dys-synchrony was estimated through a retrospective register-based study. A retrospective analysis of case reports was done wherein test findings of all those who visited the Department of Audiology, All India Institute of Speech and Hearing, Mysore between January 2000 and December 2003 were reviewed. The results showed the following:

a) prevalence of auditory dys-synchrony was 1 in 183 in individuals with permanent sensori-neural hearing loss,

b) the most common audiometric configuration seen in these patients was a 'peaked' audiogram,

c) wide variation in speech identification scores,

d) individuals with 'peaked' audiograms had better speech identification scores, and that

e) speech identification scores were significantly correlated with hearing thresholds at low frequencies. The speech identification scores decreased with decrease in thresholds at low frequencies.

1.1 Prevalence

61 of 11, 205 individuals with permanent sensory-neural hearing loss were identified to have auditory dys-synchrony which means a prevalence of 0.53% or I in 183 persons. This figure is significantly lower than the figure of 11% reported by Ranee et al. (1999) and the 1.8% reported by Tang, McPherson, Yuen, Wong, and Lee (2004). This difference might be because of the difference in the population tested. A large number of subjects in the study of Ranee et al. (1999) were bom premature and were at risk for a neurodevelopmental disorder (hyperbilirubinemia). It is known that hyperbilirubinemia can cause both permanent and temporary dysfunction of auditory pathways. Some of the infants may recover from abnormal ABR as they grow. Therefore, the figure reported by Ranee et al. (1999) seems to be an overestimate of the extent of auditory dys-synchrony. The figure of 1.8% reported by Tang et al. (2004), in school-aged children (N = 56), also seems to be an over estimation because their sample size was too small to allow any generalization. In the 56 children that they studied, only one child had evidence of auditory dys-synchrony, which meant a prevalence of 1.8%.

The results of the present study are similar to those reported (1/200) by Davis and Hirsh (1979). Actually, the prevalence of auditory dys-synchrony might be slightly higher than what is reported in the present study because some or the other test was not administered for nearly 30% of the hearing impaired in the present study, and thus, some subjects might have missed the

diagnostic criteria for auditory dys-synchrony employed in this study. Also, the present study employed transient evoked otoacoustic emissions to infer normal functioning of outer hair cells. However, Starr, Sininger and Praat (2000) have reported that around 16% of individuals with auditory dys-synchrony loose otoacoustic emissions over time. Thus, some patients with auditory dys-synchrony might have been missed out in the present study. However, irrespective of the number of patients with auditory dys-synchrony, early identification of this disorder is necessary because conventional management techniques like use of amplification devices does not seem to be effective with this population and may even be dangerous.

Only 2 of the 61 patients in the present study showed neurological deficits which is much lower than the 80% reported by Starr, Sininger and Praat. (2000). This may be because Starr et al. dealt predominately with neurological population in a medical setup. However, the All India Institute of Speech and Hearing, Mysore where the present study was conducted functions in a non-medical setup. It may be that not many patients with medical history have consulted at this institute.

Female to male ratio of auditory dys-synchrony was 2:1 in this study. This along with the report of two patients that they developed the problem after the delivery of their first child suggests the need for investigation of hormonal influence in precipitating the condition of auditory dys-synchrony.

1.2 Audiological Profile

The most common audiometric configuration seen in persons with auditory dys-syhchrony in this study was 'peaked' type, with the peak usually occurring at 2 kHz. Frequent occurrence of such an audiogram may be due to anatomico-physiological make up of the auditory nerve. The longest fibers are most susceptible to pathology. The longest cochlear nerve fibers are those going from the apex of the cochlea (which mediates low frequencies). The shortest fibers are those going to the second half of the first cochlear turn (which mediates middle frequencies). Those going to the basal parts of the cochlea have lengths in between these two extremes. These mediate high frequencies. Hence, the mid frequencies are affected less than the low and high frequencies (Starr, Picton & Kim, 2001). This is only a supposition. Our subjects were in the age range of 16-30 years, with the mean age of 22.4 years. Therefore, it is possible that high frequency hearing loss was not set in (because of young age) though they all had sensori-neural type of hearing loss. Low frequency hearing loss (500 Hz and 1 kHz) in patients with multiple sclerosis who also showed symptoms of brainstem lesions has been attributed to pathology of the central nervous system (Cohen & Rudge, 1984) and not to a lesion of the auditory nerve. However, Cohen and Rudge themselves have recognized the inadequacy of their explanation. In summary, it is unclear why mid frequencies are less affected in patients with auditory dys-synchrony in the present study.

Speech identification scores varied with audiometric configuration. Patients with 'peaked' audiograms had better speech identification scores than patients with other types of audiometric configurations. Halpin (2003) described a type of audiogram which he termed 'Tuning curve audiogram'. The characteristic of these audiograms is that they will have a single normal threshold at a given frequency with a sharp slope on either side of this frequency. Individuals with such tuning curve audiogram generally exhibit speech recognition scores that are disproportionate to their puretone thresholds. He hypothesized that such audiograms are the result of normal functioning inner hair cells located in the region subserving the peak frequency. At sufficiently high intensities, these same cells respond to other frequencies resulting in a complete audiogram. However, the 'peaked' audiogram observed in our patients is slightly different from the tuning curve audiograms described by Halpin (2003) in the sense that the low and high frequency tails of audiograms of our population was not steep. This indicates that there may be surviving inner hair cells/ functioning auditory nerve fibres at the mid frequency region while there are dead regions underlying extreme frequencies. Functioning inner hair cells/ auditory nerve fibres may be present only in the extreme frequency regions in other audiometric configurations. For example, at 8 kHz in raising, at 250 Hz and 500 Hz in sloping, and at 250 Hz and at 8 kHz in saucer shaped configuration. Mid-frequency bands contribute more to speech intelligibility than lower or higher frequencies. Presence of dead region in the cochlea/auditory nerve fibers at mid frequencies might be responsible for poor speech identification scores, in addition to temporal

asynchrony in auditory nerve firings, in patients with rising, sloping and saucer-shaped audiograms than in patients with peaked audiograms. Even though there was a significant negative correlation between thresholds at all frequencies and speech identification scores except at 8 kHz, scatter plots showed no specific trend between the two variables except at low frequencies. No relation between speech identification scores and behavioral thresholds are expected either as inability of theses patients to understand speech is because of temporal asynchrony in auditory system and not due to elevated thresholds, per se.

This frequency specific correlation may be related to the underlying physiology of coding of high frequencies and low frequencies. Low frequencies are usually coded by phase locked responses in type I auditory nerve fibers. Individuals with auditory dys-synchrony cannot use phase locking cues to the same extent as normal hearing listeners due to dyssynchronous firing of auditory nerve fibers. However, detection of high frequencies does not depend on phase locking cues as much as low frequencies do. On the other hand, it depends on information on the place of excitation on the basilar membrane. We hypothesize that hearing sensitivity at low frequencies in these individuals may indicate the extent of temporal disruption in the auditory system. Therefore, the greater the loss at low frequencies, the greater is the severity of temporal asynchrony which in turn reduces speech perception abilities.

Mean amplitude of TEOAE for adults in our clinic is 11.5 dB SPL. The amplitude of TEOAE in individuals with auditory dys-synchrony in this

study was 16 dB SPL. Higher amplitude of evoked otoacoustic emission in patients with auditory dys-synchrony has also been reported by others (Hood & Berlin, 2001). This phenomenon has been attributed to lack of efferent suppression of otoacoustic emissions. As there is evidence to say that the efferent system helps in the perception of speech (Kumar & Vanaja, 2004; Muchink, et al., 2004), and high amplitudes of TEOAE reflects the non-functioning efferent system, one would expect poor speech identification scores to be correlated with high amplitudes of TEOAE. However, this was not observed in the present study. Amplitude of TEOAE also did not correlate with behavioral threshold at any of the frequencies. Hence, OAEs were not helpful in predicting speech identification scores or behavioral thresholds in patients with auditory dys-synchrony.

Starr, Sininger and Praat (2000) have reported that around 16% of individuals with auditory dys-synchrony loose otoacoustic emissions over time. This means that some of the auditory dys-synchronics cannot be identified from OAE testing. Though cochlear microphonics are present in a majority of auditory dys-synchronics it is difficult to record cochlear microphonics in an out patient ward. Also, as said earlier, there was no significant correlation observed between amplitude of OAE and speech identification scores. One would expect a significant correlation between these two as higher amplitude of OAE reflects nonfunctioning efferent system and efferent system helps in perception of speech. This casts serious doubts on the utility of OAE testing in the study of auditory dys-synchrony except to verify the integrity of outer hair cells. On the other hand, all individuals tested

in the present study demonstrated significantly poorer JNDs and modulation detection thresholds. These results show that testing for the perception of temporal parameters is a more efficient and effective means of screening for auditory dys-synchrony than testing for OAEs.

2.0 Temporal Parameters Influencing Speech Perception in Persons with Auditory Dys-Syncbrony

Perception of three temporal parameters of speech, namely, transition duration (TD), burst duration (BD) and voice onset time (VOT) was investigated in the second phase of the study. Some of the important findings of experiments in this phase are as follows:

a) individuals with auditory dys-synchrony demonstrated elevated JNDs for TD, BD and VOT indicating deficient processing of temporal information

b) lengthening of TD improved perception of both place and manner information,

c) lengthening of VOT resulted in improved perception of voicing feature, and

d) speech stimuli which were modified to reflect best JNDs for TD, BD and VOT did not significantly improve speech perception compared to stimuli which incorporated lengthened TD or BD or VOT.

Temporal processing deficits in individuals with auditory dys-synchrony have been well documented using non-speech signals (Rance,

Mckay, & Grayden, 2004; Zeng, Oba, Garde, Swinger, & Starr, 1999; Zeng, Kong, Michalewsk, & Starr, 2005). Our data indicated that individuals with auditory dys-synchrony had difficulty in differentiating CV syllables that were modified in respect of TD, BD and VOT. The JNDs of subjects with auditory dys-synchrony were almost three to four times larger than those for normal hearing listeners. Elevated JNDs for these temporal parameters of speech mean that auditory dys-synchronics have difficulty in discriminating speech sounds that differ in temporal aspects. Kraus et al. (2000) reported that a subject with auditory dys-synchrony had marked difficulty in discriminating speech sounds that differed in spectral onset. Transitions and bursts are rapid spectro-temporal changes at the onset of stimulus. These two observations together mean that individuals with auditory dys-synchrony have difficulty in processing temporal and spectral information at the stimulus onset.

Confusion matrices and SINFA showed that lengthening of TD resulted in maximum information transfer, followed by lengthening of VOT and BD, in that order. Results of SINFA should be interpreted with caution. SINFA is robust when the stimuli occur with equal frequency. But, in the present series of experiments, the number of stimuli was slightly different for different speech sounds depending on whether or not they were identified in the experimental condition of presentation of unmodified stimuli. Obviously, speech syllables which were identified in unmodified condition were not included in subsequent experiments. This limitation should be kept in mind while interpreting the results of SINFA. It has been shown that individuals with auditory dys-synchrony have difficulty in processing short duration

sounds (Zeng et al., 2005). Formant transitions are short (< 50 ms), but, they are important acoustic cues for place of articulation in stop-vowel syllables. Lengthening of transition duration may have enhanced their utility for speech perception, but how of this process is not clear. Similar results with stimuli with lengthened transition duration have been reported by other investigators in language learning impaired children (Tallal & Piercy, 1975; Tallal et al., 1996). Tallal and her colleagues showed that speech perception in language learning impaired children improved when the duration of formant transitions and modulation depth were increased. We hypothesize that lengthening of transition duration reduces modulation frequency without altering the modulation depth and overall spectrum of the stimulus. This hypothesis was tested by lengthening sinusoidally amplitude modulated white noise by a factor of two. Lengthening resulted in a reduction of modulation frequency, but the modulation depth and spectrum were unaltered. Figure 5.1 shows the modulation frequency, modulation depth and spectrum of the modified (a) lengthened (b) white noise and (c) the spectrum. Reduction in modulation frequency augments speech perception in individuals with auditory dys-synchrony as their modulation detection is better at low modulation frequencies than at high modulation frequencies and there is good correlation between the perception of modulation depth and speech perception abilities.

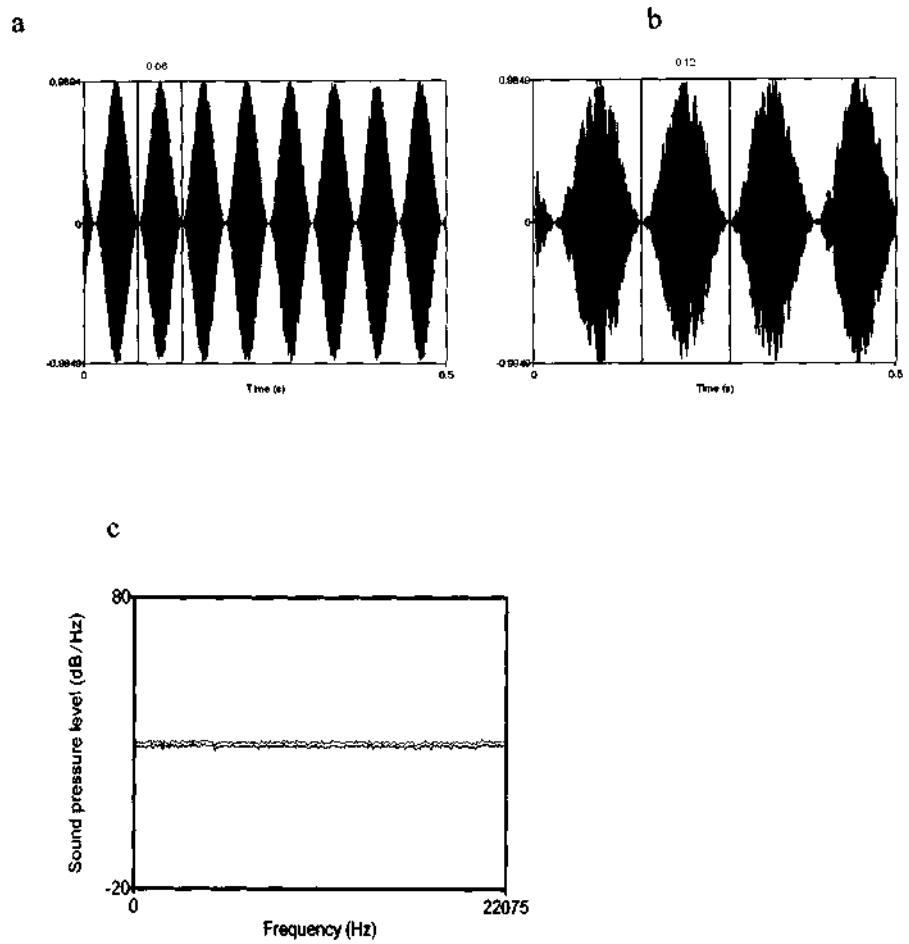


Figure 5.1 a) Sinusoidally amplitude modulated noise at 16 Hz. b) The same sinusoidally amplitude modulated noise lengthened two times. Note that modulation frequency is reduced, yet modulation depth remains the same, c) Spectrum of unmodified (black line) and lengthened (red line) sinusoidally amplitude modulated white noise.

Lengthening of transition duration improved perception of both place and voicing information. Lengthening of VOT resulted in greater improvement in voicing cue than place cue. One of the factors that contributes to poor speech perception in persons with auditory dys-synchrony is elevated

gap detection threshold (Zeng et al., 1999). Gap detection ability has been associated with discrimination of syllables that vary in voice onset time (Dreschler & Plomp, 1985; Tyler, Summerfield, Wood, & Femandes, 1982). It may be inferred from these observations that some of the difficulties faced by individuals with auditory dys-synchrony in the perception of voicing may be due to their inability to perceive short duration VOT. Lengthening of VOT might have made this acoustic cue more salient and helped persons with auditory dys-synchrony to perceive voicing. Furthermore, individuals with auditory dys-synchrony have difficulty in processing short duration signals, more so in the presence of noise. Zeng et al. (2005) have demonstrated forward and backward masking effects that are 3-4 times larger in subjects with auditory dys-synchrony than in normal listeners. These observations, in combination, suggest that the vowel portion following/preceding the VOT might mask (forward or backward as the case may be) the VOT and pose difficulty in perceiving this voicing cue. It is plausible that increasing the duration of VOT would reduce masking effects and which in turn facilitates better processing of VOT in persons with auditory dys-synchrony. Even then this result was not as per our prediction. We hypothesized that voicing (VOT) being such a distinct physical phenomenon, lengthening it would enhance speech perception. Though lengthened VOT did improve identification of voicing and thus speech perception, there were other short elements of speech like TD whose lengthening promoted higher transfer of information (better speech identification score). This defies any explanation.

Stimuli used in the experiments of Phase II were uttered by a male speaker. Since male voice is rich in low frequency and individuals with auditory dys-synchrony have difficulty in processing low frequencies, it would be interesting to speculate on the perception of the speech of a female speaker. It may or may not be different. Future research can look into this aspect.

2.1 Modulation Detection in Persons with Auditory Dys-synchrony

Modulation detection thresholds for sinusoidally amplitude modulated white noise at 4 Hz, 16 Hz, 32 Hz, 64 Hz, 128 Hz and 200 Hz were measured in Phase III of the study. Previous studies have reported significant temporal processing impairment, and a significant correlation between impairment of temporal processing and deficits in speech perception in individuals with auditory dys-synchrony (Ranee et al, 2004; Zeng et al, 2005; Zeng et al, 1999). Similar results were obtained in the present study. Peak modulation detection thresholds were significantly correlated with speech identification scores obtained for unmodified stimuli, for speech stimuli modified for TD as well as for speech stimuli modified for TD, BD and VOT (combined). The peak modulation detection thresholds in individuals with auditory dys-synchrony were about 3 times higher than those in normal hearing listeners. Individuals with auditory dys-synchrony had more problems in processing high (> 128 Hz) modulation frequency than low. Inability of many subjects to perceive amplitude modulation of 0 dB (100%) at modulation frequencies

higher than 128 Hz indicates the importance of temporal synchrony in auditory perception.

We propose a phenomenological model to explain poor perception of temporal modulations in individuals with auditory dys-synchrony. This model is based on the model of auditory dys-synchrony proposed by Zeng et al (2005) as well as information available on processing of amplitude modulated sounds in the auditory nerve and lower brainstem (see Frisina, 2001; Jons, Schreiner & Rees, 2004, for a detailed discussion). Auditory nerve fibers and onset units of cochlear nuclei are highly synchronized and phase locked to the envelope of amplitude modulated signals. It has also been demonstrated that different sets of cochlear neurons are tuned to different amplitude modulation frequencies. The synchronization or phase locking of the discharges of the auditory nerve fibers as well as cochlear neuron units to the envelope of amplitude modulation reduces with increase in modulation frequency (see Frisina, 2001, Figure 5, pp 7). Figure 5.2 shows synchronized firing of three auditory nerve fibers (a, b & c). Trace (d) represents the combined output response of auditory nerve fibers or cochlear nuclei (average of a, b & c). The bottom trace (e) represents the unmodulated and sinusoidally amplitude modulated stimulus. Neural synchrony preserves modulation in terms of temporal discharges relative to unmodulated noise and background spontaneous random neural activity. Figure 5.3 shows responses to high modulation frequency. Figure 5.4 shows the model based on dys-synchronized nerve conduction in three demyelinated nerve fibers at low

modulation frequency. Due to dys-synchrony, these fibres respond differentially to modulation, and the result is a smeared representation of the stimulus. Even though the average neuronal output is smeared, some amount of temporal information is still preserved which helps in detecting modulation at high modulation depths (as the modulation depth increases, synchronization also improves). Figure 5.5 shows neural responses at high modulation frequencies. Due to dys-synchronous firing of the nerve fibers, the average response is smeared and it completely abolishes the coding of modulation information. Average responses to high modulation frequencies appear similar to those for unmodulated stimulus (Figure 5.5d).

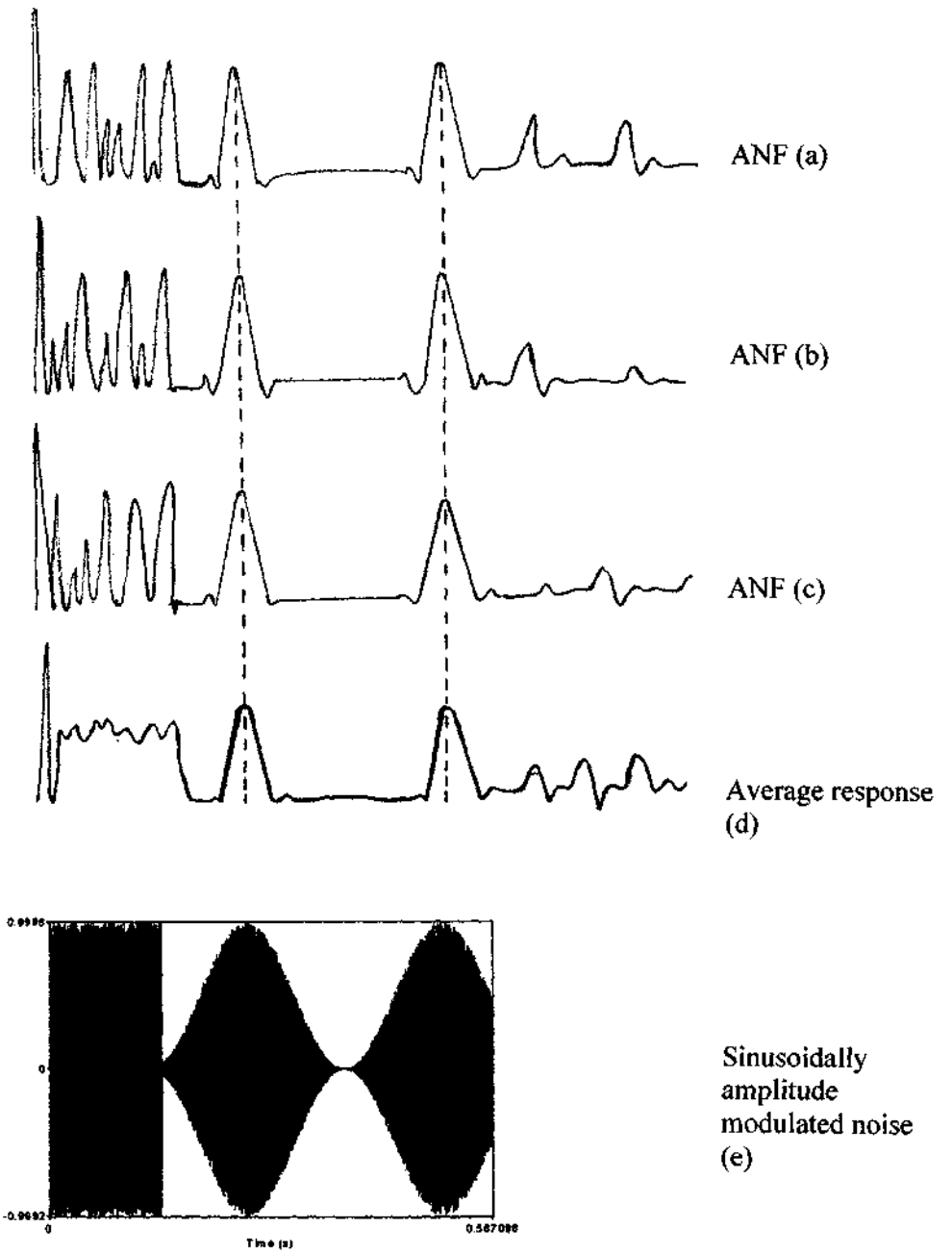


Figure 5.2. Synchronous responses of 3 auditory nerve fibers (ANF) to amplitude modulated noise (low modulation frequency).

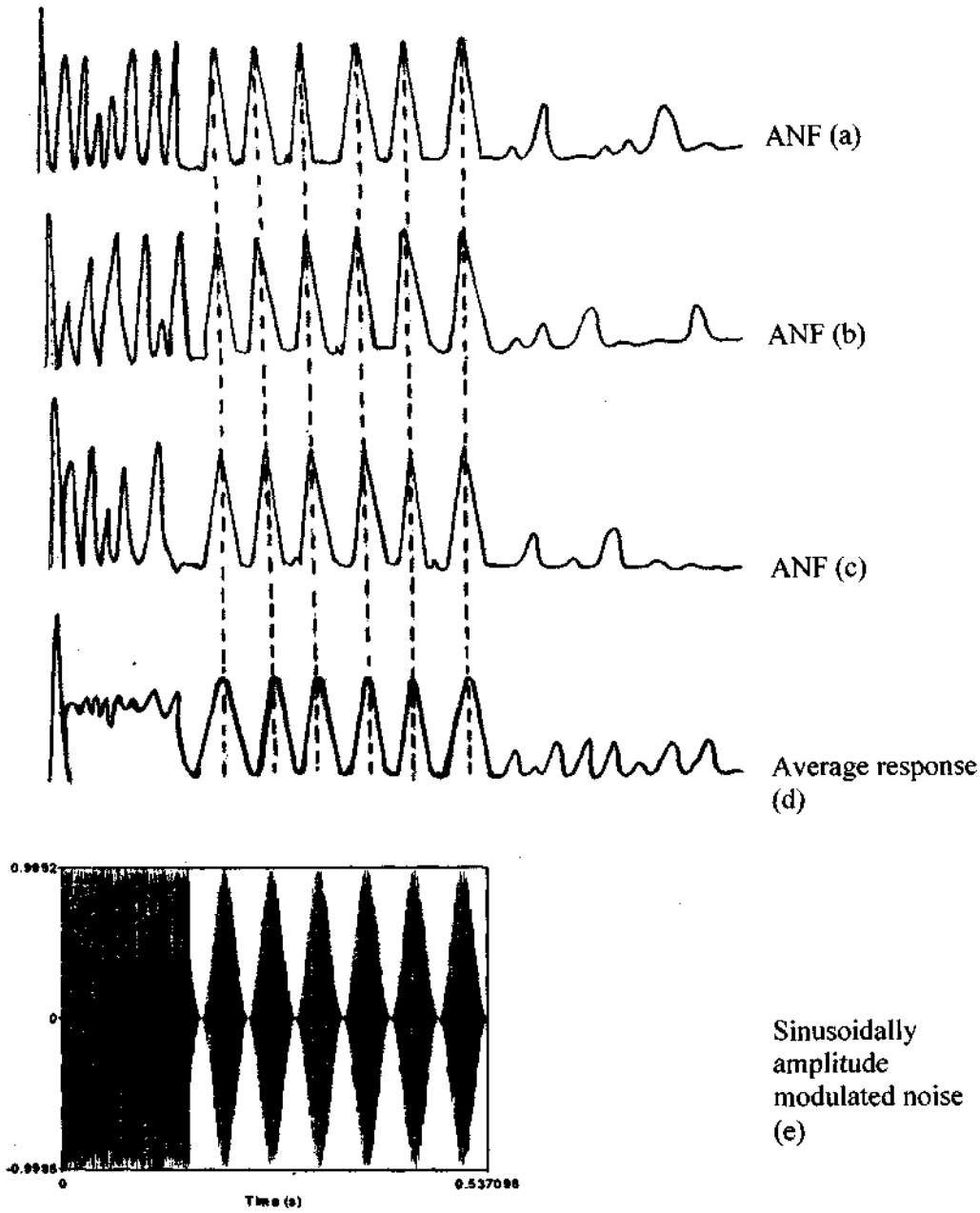


Figure 5.3. Synchronous responses of 3 auditory nerve fibers (ANF) to amplitude modulated noise (high modulation frequency).

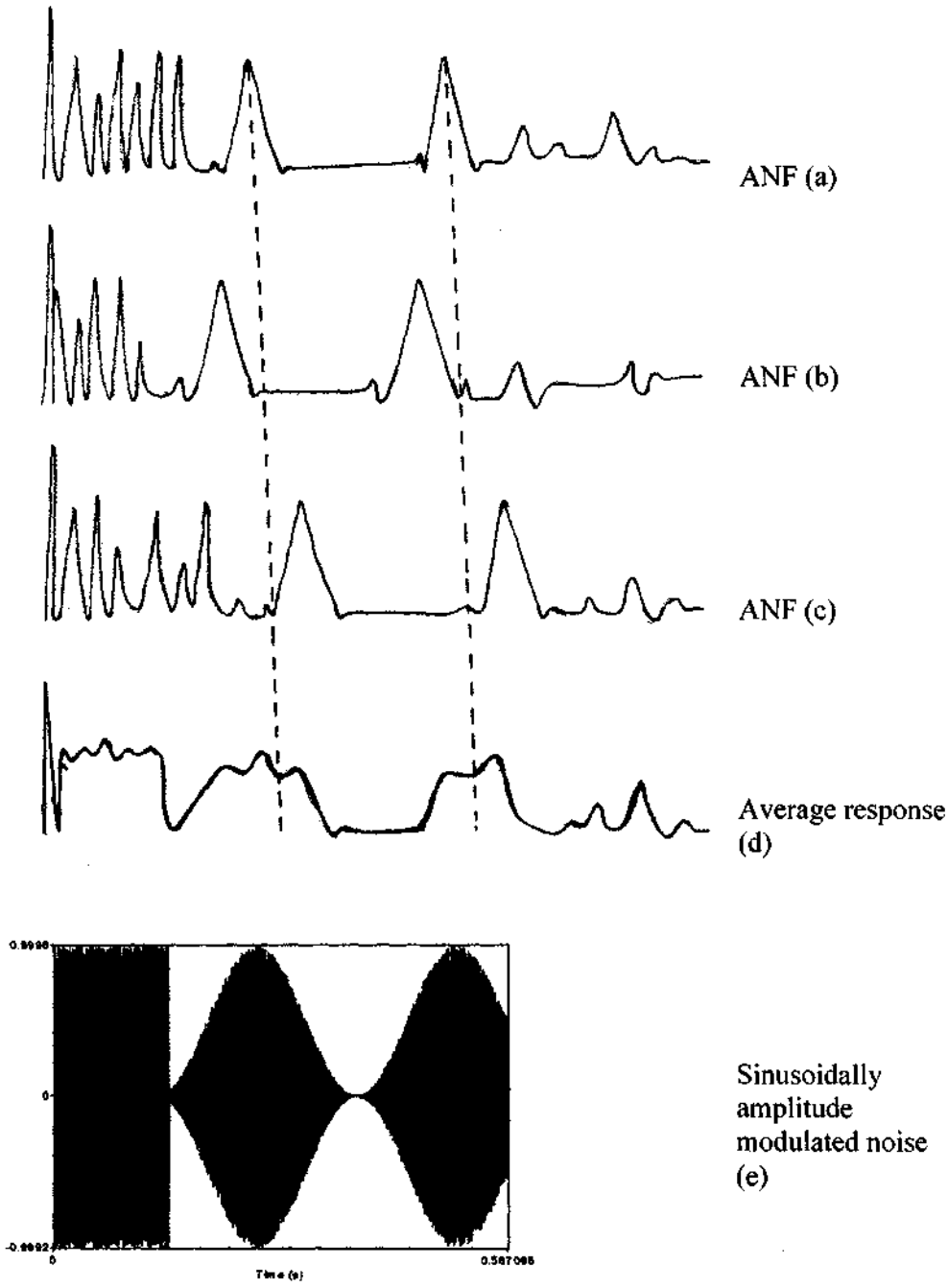


Figure 5.4. Dys-synchronous responses of 3 auditory nerve fibers (ANF) to amplitude modulated noise (low modulation frequency).

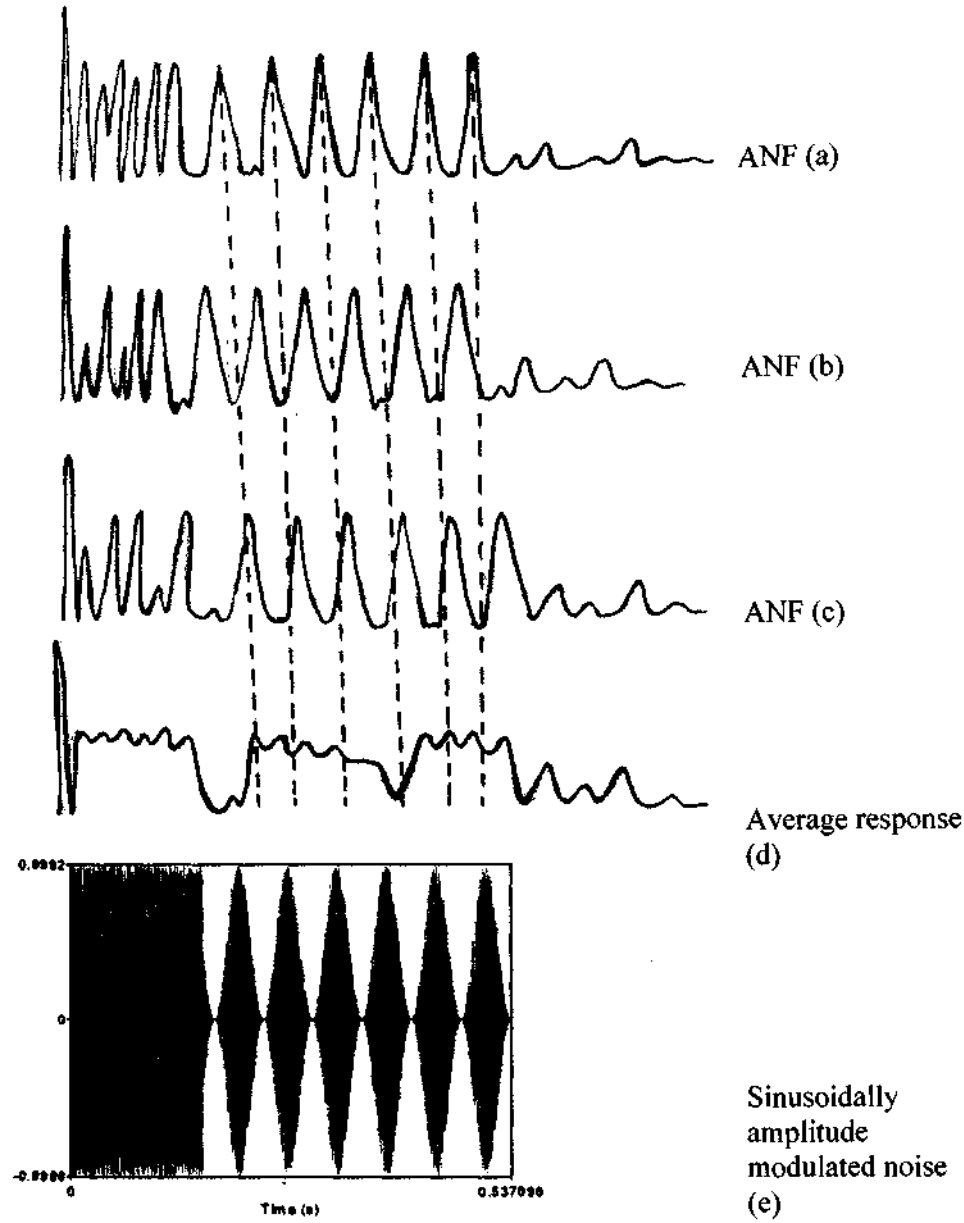


Figure 5.5. Synchronous responses of 3 auditory nerve fibers (ANF) to amplitude modulated noise (high modulation frequency).

Effects of reduced temporal fluctuations on speech perception in normal listeners have been reported previously (Drullman, Festen & Plomp, 1994). Elevated modulation detection threshold at low modulation frequencies in combination with virtually complete absence of perception of modulations at high modulation frequencies is sufficient to disrupt perception of amplitude envelope cues in normal speech. In the auditory system, higher modulation frequencies are processed at the level of auditory nerve and brainstem, whereas lower modulation frequencies are processed mainly in the thalamus and auditory cortex. As one ascends the auditory system, a shift in neural encoding occurs. Synchronous response for temporal coding is emphasized at the level of auditory nerve and brainstem (which codes high modulation frequencies) while less and less emphasis is placed on synchrony as one moves up centrally (which code low modulation frequencies - Frisina, 2001). This may be the reason for the greater problem that persons with auditory dys-synchrony have in processing high rate of modulations, which require synchronous firing of auditory nerve fibers in comparison with processing of low frequency modulations. However, only peak sensitivity was considered in all the analyses in this study. Reduced peak sensitivity may also be due to reduced ability of auditory dys-synchronics to perceive amplitude changes. But, this is least likely because presentation level was changed in all the subjects at least at one modulation frequency and modulation detection threshold was rechecked to ensure that subjects were not using loudness judgments.

At present, cochlear implant is the most preferred treatment of choice for individuals with auditory dys-synchrony. But, the degree of hearing loss may not justify use of cochlear implant in some patients with auditory dys-synchrony. Our results suggest that lengthening of transition duration of stop consonants does improve speech identification scores in individuals with auditory dys-synchrony. The results on speech perception from the present study suggest, in combination with data from previous psychophysical research (Zeng et al., 1999, 2005), that innovative processing strategies that lengthen transition duration of speech sounds in a hearing aid will bring in significant benefits for individuals with auditory dys-synchrony. Other schemes that may help in facilitating speech perception in individuals with auditory dys-synchrony include lengthening of other short elements of speech such as voice onset time. It has been shown that individuals with auditory dys-synchrony have problems in perceiving short duration acoustic stimuli, but not long (Zeng et al., 2005), and thus lengthening these short elements may help auditory dys-synchronics in their perception of speech.

An interesting observation from the present study is with respect to the age range of auditory dys-synchronics in this study and age of on set of the problem in them. Subjects in the second phase of the study were in the age range of 16-30 years (mean age 22.4 years) when their problem was identified. Sininger and Oba (2001) have reported that the problem of auditory dys-synchrony had been identified in 75% of their patients before they were 10 years, and nearly 40% had been identified before they were two years of age. This means that problem is identified at a much later age

in this country thus depriving such subjects, particularly children, from valuable auditory stimulation during their language acquisition years. There is no need to reiterate that such difficulty in speech perception will have consequences for the child's overall development, particularly educational. Therefore, there is a great need to put in a mechanism to identify such children with auditory dys-synchrony at the earliest age possible.

CHAPTER 6

SUMMARY AND CONCLUSIONS

The purpose of this study was to (a) estimate the prevalence of auditory dys-synchrony in Mysore, a city of one million population in Southern India, (b) evaluate perception of temporal parameters of speech in individuals with auditory dys-synchrony, and (c) measure modulation detection thresholds for sinusoidally amplitude modulated white noise in individuals with auditory dys-synchrony.

The study was conducted in three phases. In Phase I, prevalence of auditory dys-synchrony in Mysore city was estimated using a register-based study. Data pertaining to nearly 21,500 patients with hearing impairment who visited All India Institute of Speech and Hearing, Mysore between Jaunary 2000 and December 2003 was analysed for prevalence of auditory dys-synchrony.

Perception of three temporal parameters of speech, namely, transition duration (TD), burst duration (BD) and voice onset time (VOT) were investigated for their influence on speech perception in individuals with auditory dys-synchrony in Phase II of the study. Four experiments were designed for this purpose: Speech identification scores for unmodified stimuli were measured in 30 auditory dys-synchronics in Experiment I. The second experiment was on determining just noticeable difference (JND) for TD, BD and VOT in CV (consonant + vowel) syllables. CV syllables with voiceless stop consonants - velar /k/, alveolar /t/, retroflex /t/, and bilabial /p/ - and their

voiced cognates were used in the study. JND was determined using an adaptive tracking technique - Parameter Estimation through Sequential Testing (PEST) with an "AX same-difference" discrimination paradigm.

The effect of lengthening of each of these temporal parameters on speech perception in individuals with auditory dys-synchrony was evaluated in the third experiment. The fourth experiment was designed to evaluate perception of speech in individuals with auditory dys-synchrony for CV stimuli which had JNDs which had yielded the maximum speech identification score in Experiment 3.

Modulation detection thresholds for sinusoidally amplitude modulated white noise at 4 Hz, 16 Hz, 32 Hz, 64 Hz, 128 Hz and 200 Hz were measured in Phase HI of the study.

Some of the important findings of this study can be summarized as follows:

a) Prevalence of auditory dys-synchrony among adults with sensory- neural hearing loss in India is about 0.53% or 1 in 183. It is only 0.28% (1 in 348) if all persons with hearing impairment are considered.

b) 46% of the individuals with auditory dys-synchrony had peaked audiograms and individuals with peaked audiograms had significantly better speech identification scores than individuals with other types audiometric configuration.

c) Puretone hearing thresholds at low frequencies were significantly correlated with speech identification scores.

d) JNDs for TD, BD and VOT were significantly poorer in individuals with auditory dys-synchrony than in normal hearing subjects.

e) Lengthening of TD significantly improved speech identification scores in auditory dys-synchronics than lengthening of BD and VOT.

f) 'Sequential information transfer analysis (SINFA)' showed that lengthening of TD resulted better in transmission of place and voice features.

g) Combined modification of TD, BD and VOT did not bring about significant improvement in speech identification scores compared to lengthening of TD alone.

h) Peak modulation detection thresholds were significantly lower in individuals with auditory dys-synchrony compared to normals.

i) Peak modulation detection thresholds were significantly correlated with speech identification scores obtained for stimuli modified for TD, combined stimuli (modified for TD, BD, VOT) and unmodified stimuli.

These results have been interpreted in the light of findings from psychophysical as well as neurophysiological research on hearing. Poor temporal processing in persons with auditory dys-synchrony was explained with a phenomenological model. Lengthening of TD resulted in significant improvement in speech identification scores in individuals with auditory dys-synchrony. Lengthening of TD resulted in 100% speech identification scores in five of the 30 subjects. Hearing aids that incorporate algorithms to lengthen transition in conversational speech may prove beneficial for individuals with

auditory dys-synchrony. Poor modulation detection thresholds in individuals with auditory dys-synchrony reflect their inability to process amplitude envelope cues. Therefore, we suggest that amplitude compression in hearing aids should be avoided for individuals with auditory dys-synchrony.

Though, at present, cochlear implantation is the preferred treatment option for auditory dys-synchronics, our behavioural data suggests that signal processing strategies that enhance temporal information will benefit these individuals. One such option is to lengthen TD and VOT. Other potential strategies that may improve speech perception in individuals with auditory dys-synchrony include (a) lengthening of other short duration temporal cues such closure duration, syllable duration, (b) enhancement of modulations in the formant transition regions, (c) transposition of higher modulation frequencies to low modulation frequency as individuals with auditory dys-synchrony have severe problems in perceiving high frequency modulations. These features are not present in today's hearing aids. Efforts are on way at our laboratory to incorporate some of these modifications in digital hearing aids and see if they are indeed beneficial to auditory dys-synchronics.

The significance of this study should be seen in the following context: there are as many studies which say that hearing aids are useful to subjects with auditory dys-synchrony as those which have failed to show these beneficial effects. The present study provides indirect evidence that hearing aids which process speech to enhance certain critical short events may be beneficial to persons with auditory dys-synchrony. One such factor is enhancement (lengthening) of transition duration which seems to improve

perception of both manner and place cues. VOT is another factor whose lengthening seems to benefit auditory dys-synchronics in their perception of speech.

This study also throws light, for the first time, on the extent of the problem in India. Prevalence of 1 in 183 means that the problem occurs much more frequently than believed all these years. Data such as this based on a completely random sample should be considered intrinsically very strong. The somewhat high (1/183) prevalence together with the controversy on the benefit that the auditory dys-synchronics derive from amplification means that management of this problem is quite a challenge in our country. It is also hoped that the high prevalence of auditory dys-synchrony would draw the attention of planers and public to this problem, particularly in Indian context, and help in initiating discussion on aids and concessions to be given to this population on par with individuals with other disabilities.

REFERENCES

Apoux, F., Tribut, N., Dehruille, X., & Lorenizie, C. (2004). Identification of envelope expanded sentences in normal hearing and hearing impaired listeners. Hearing Research, 189, 13-24.

Apoux, F., Crouzet, O., Lorenzi, C. (2001). Temporal envelop expansion of speech in noise for normal hearing and hearing impaired listeners: Effects on identification performance and response times. Hearing Research, 153, 123-131.

Berlin, C. I. (1999). Auditory Neuropathy: Using OAEs and ABRs from screening to management. Seminars in Hearing, 20, 307-315.

Berlin, C. I., Hood, L. J., Hurely, A., & Wen, H. (1996). Hearing aids: Only for hearing impaired patients with abnormal otoacoustic emissions. In C.I. Berlin (Ed.), Hair cells and hearing aids (pp 99-111). San Diego: Singular publishing group.

Berlin, C, Hood, L., Morlet, T., Rose, K., & Brashears, S. (2003). Auditory Neuropathy/Dys-synchrony: Diagnosis and Management. Mental Retardation and Development Disabilities Research Reviews, 9, 225-231.

Berlin, C.I, Morlet, T., & Hood, L. (2003). Auditory neuropathy/dyssynchrony: Its diagnosis and management. The Pediatric Clinics of North America, JO, 331-340.

Berlin, C.I., Bordelon, J., St John, P., Wilensky, D., Hurleym, A., Kluka, E., & Hood, L.J. (1998). Reversing click polarity may uncover auditory neuropathy in infants. Ear and Hearing, 79, 37-47.

Berlin, C.I., Hood, L., & Rose, K. (2001). On renarning auditory neuropathy as auditory dys-synchrony. Audiology Today 13 (6) 15-17.

Berlin, C.I., Hood, L.J., Morlet, T., Li, L., Brashears, S., Tedesco, S., Rose, K., Wilensky, D., Jeanfreus, J. (2003). Auditory neuropathy/auditory dys-synchrony (AN/AD): Management and results in 193 patients [Abstract]. Association for research in Otolaryngology, 191.

Blackburn, C.C., & Sachs, M.B. (1989). Classification of units types in the anteroventral cochlear nucleus: PST histograms and regularity analysis. Journal of Neurophysiology, 62, 1303-1329.

Bradlow, A.R., Kraus, N., Nicol, T.G., Mcgee, T.J., & Cunningham, J. (1999). Effects of lengthened formant transition duration on discrimination and neural representation of synthetic CV syllables by normal and learning-disabled children. Jouranl of the Acoustical Society of America, 106, 2086-2096.

Buss, E., Labadie, R.F., Brown, C.J., Gross, A.J., Grose, J.H., & Pillsbury, H.C. (2002). Outcome of cochlear implantation in pediatric auditory neuropathy. Otology and Neurotology, 23, 328-332.

Cohen, M., & Rudge, P. (1984). Effect of multiple sclerosis on puretone thresholds. Acta Otolaryngology, 97, 291-295.

Cunningham, J., Nicol, T., King, C, Zecker, S.G., & Kraus, N. (2002). Effects of noise and cue enhancement on neural responses to speech in auditory midbrain, thalamus and cortex. Hearing Research, 169, 97-111.

Dallos, P., & Corey, M.E. (1991). The role of outer hair cell motility in cochlear tuning. Current Opinion in Neurobiology, 1, 215-220.

Davis, H., & Hirsh, S.K. (1979). A slow brainstem response for low frequency audiometry. Audiology, 18, 445-465.

DeFillippo, C.L., & Snell, K. (1986). Detection of temporal gaps in low frequency narrow band signals by normal hearing and hearing impaired listeners. Journal of the Acoustical Society of America, 80, 1354-1358.

Deltenre, P., Mansbach, A.L., Bozet, C, Christiaens, P., Barthelemy, P., Paulissen, D., Renglet,T. (1998). Auditory neuropathy with preserved cochlear microphonics and secondary loss of otoacousn'c emissions. Audiology, 38, 187-195.

Dreschler, W.A., & Plomp, R. (1985). Relation between psychophysical data and speech perception for hearing impaired subjects II. Journal of the Acoustical Society of America, 78, 1261-1270.

Drullman, R., Festen, J.M., & Plomp, R. (1994). Effect of temporal envelope smearing on speech perception. Journal of the Acoustical Society America, 95, 1053-1064.

Duan, J., & Wang, J. (2002). The ECochG in patients with auditory neuropathy. Lin Chuang Er Bi Yan Hou Ke Za Zhi, 605-606.

Erber, N.P. (1972). Speech envelope cues as an acoustic aid to lip reading for profoundly deaf children. .Journal of the Acoustical Society of America, 51, 1224-1227.

Frisina, R.D. (2001). Subcortical neural coding mechanisms for auditory temporal processing. Hearing Research 158, 49-54.

Glasberg, B.R., & Moore, B.C.J. (1989). Psychoacoustic ability of subjects with unilateral and bilateral cochlear impairments and their relationship to the ability to understand speech. Scandinavian Audiology, 32, 1-25.

Goldberg, J.L., & Brownell, W.E. (1973). Discharge characteristics of neurons in anteroventral and dorsal cochlear nuclei of cat. Brain Research, 64, 35-54.

Gordon- Salant, S., & Fitzgibbons, P.J. (1993). Temporal factors and speech recognition performance in young and elderly listeners. Journal of Speech and Hearing Research, 36,1276-1285.

Gordon-Salant, S. (1986). Recognition of natural and time intensity altered CVs by young and elderly subjects with normal hearing. Journal of the Acoustical Society of America, 80,1599-1607.

Gupta, A.K., Raj, H., & Anand, N.K. (1990). Auditory brainstem responses in neonates with hyperbilirubinemia. Indian Journal of Paediatrics, 58, 849-855.

Halpin, C. (2003). Tuning curve audiograms. American Journal of Audiology, 11, 56-64.

Hood, L.J., & Berlin, C.I. (2001). Auditory neuropathy/ (auditory dys-synchrony) disables efferent suppression of otoacoustic emissions. In Y. Sininger & A. Starr (Eds.), Auditory neuropathy: A new perspective on hearing disorder (pp. 183-202). Canada: Singular publishing group.

Huffman, R.F., & Henson, O.W.J. (1990). The descending auditory pathways and acoustic motor systems: Connection with inferior colliculus. Brain Research Review, 15, 295-323.

Jayaram, M. (1984). Distribution of stuttering in sentences: Relationship to sentence length and clause position. Journal of Speech Hearing Research, 27, 38-42.

Joris, P. X., Schreiner., & Rees. (2004). Neural processing of amplitude-modulated sounds. Physiology review, 84, 541-577.

Keating, P., & Blumstein, S.E. (1978). Effects of transition length on perception of stop consonants. Journal of the Acoustical Society of America, 64, 57-64.

Kraus, N. (2001). Auditory neuropathy: An historical and current perspective. In Y. Sininger, & A. Starr (Eds.), Auditory neuropathy: A new perspective on hearing disorder (pp. 1-14). Canada: Singular publishing group.

Kraus, N., Bradlow, A.R., Cheatham, M.A., Cunningham, J., King, C D., & Koch, C D. (2000). Consequences of neural asynchrony: A case of auditory neuropathy. Journal of Association for Research in Otolaryngology, 1, 33-45.

Kraus, N., Ozdamar O., Stein L., & Reed N. (1984). Absent auditory brainstem response: Peripheral hearing loss or brainstem dysfunction. Laryngoscope, 94, 400-406.

Kruase, J.C., & Braida, L.D. (2004). Acoustical properties of naturally produced clear speech at normal speaking rates. The Journal of the Acoustical Society of America, 115, 362-378.

Krishna, G. (2001). Dichotic CV-test. Revised normative data for children. Unpublished independent project, University of Mysore, Mysore, India.

Kumar, U.A., & Jayaram, M. (2005). Processing of modulation through a digital hearing aid. Manuscript submitted.

Kumar, U.A. & Vanaja, C.S. (2004). Functioning of olivocochlear bundle and speech perception in noise. Ear and Hearing, 25, 142-146.

Lacerda, F. (2001). Re-assessing perceptual consequences of CV-transition speed. Lund University, Department of Linguistics, Working papers, 49, 98-101.

Li, F., Wang, H., Chen, J., & Liang, R. (2005). Auditory neuropathy in children (Analysis of 14 cases),Lin chuanger bi yan hou ke za zhi 19, 19-21.

Lorrenzi, C, Berthommier, F., Apoux , F., & Bacri, N. (1999). Effects of envelope expansion on speech recognition, Hearing Research, 136, 131-138.

Lynne, A., Werner, & Gray, L. (1998). Behavioral studies of hearing development In E. Rubel, A. Popper & R.R. Fay (Eds.), Development of auditory system (pp. 15-47). NY: Springer.

Mchalewski, H. J., Stair, A., NguyenKong, T.T., & Zeng, F.G. (2005). Auditory temporal process in normal hearing individuals and in patients with auditory neuropathy. Clinical Neurophsiology 116, 669-680.

Meizenich, M., Jenkins, W.M., Johnston, P., Schreiner, C, Miller, S., & Tallal, P. (1996). Temporal processing deficits of language learning impaired children ameliorated by training. Science, 277, 77-81.

Miyamoto, R.T., Kirk, K.I., Renshaw, J., & Hussain, D. (1999). Cochlear implantation in auditory neuropathy. laryngoscope, 109, 181-185.

Moore, B.C.J. (2003). An introduction to the psychology of hearing (5th Ed.) San Diego: Academic Press.

Moulines, E., & Lorche, J. (1995). Non parametric techniques for pitch scale and time scale modification of speech. Speech Communication, 76, 175-205.

Muchink, C, Roth, D.A.V., Jebara, R.O., Katz, H.P., Shabtai, E.L, & Hildesheimer, M. (2004). Reduced medial olivocochlear bundle system function in children with auditory processing disorders. Audiology and Neurootology, 9, 107-114.

Peterson, A., Shallop, J., Driscoll, C, Breneman, A., Babb, J., Stoeckel, R., Fabry, L. (2003). Outcomes of cochlear implantation in children with auditory neuropathy. Journal of American Academy of Audiology, 188-201.

Picheny, M.A., Durlach, N., & Braida, L.D. (1985). Speaking clearly for the hard of hearing I: Intelligibility difference between clear and conversational speech. Journal of Speech and Hearing Research, 2&, 96-103.

Picheny, M.A., Durlach, N., & Braida, L.D. (1986). Speaking clearly for the hard of hearing. II: Acoustical characteristics of clear and conversational speech, Journal of Speech and Hearing Research, 29, 434-446.

Plomp, R. (1988). The negative effect of amplitude compression in multichannel hearing aids in the light of the modulation-transfer function. Journal of the Acoustical Society of America, 83, 2322-2327.

Price, P.J. & Simon, H.J. (1984). Perception of temporal differences in speech by normal hearing adults: Effects of age and intensity. Journal of the Acoustical Society of America, 72, 405-410.

Ranee, G., Beer, D.E., Cone-Wesson, B., Shepherd, R.K.., Dowell, R.C., King, A.M., Rickards, F.W., & Clark, G.M. (1999). Clinical findings for a group of infants and young children with auditory neuropathy, Ear and Hearing 20, 238-252.

Ranee, G., Cone-Wesson, B., Wunderlich, J., & Dowell, R. (2002). Speech perception and cortical event related potentials in children with auditory neuropathy. Ear and Hearing, 23, 239-253.

Ranee, G., McKay, C, & Grayden, D. (2004). Perceptual characterization of children with auditory neuropathy. Ear and Hearing, 25, 34-46.

Santarelli, R., & Arslan, E. (2002). Electrocochleography in auditory neuropathy. Hearing Research, 170, 32-47.

Sek, A., & Moore, B. C. (1995). Frequency discrimination as a function of frequency, measured in several ways. Journal of the Acoustical Society of America, 97, 2479-2486.

Shallop, J.K., Peterson, A., Facer, G. W., Fabry, L.B., & Driscoll, C. L. (2001). Cochlear implants in five cases of auditory neuropathy: post operative findings and progress. Laryngoscope, 111, 555-562.

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

Shivprakash, K. (2003). Gap detection test-Development of normative data. Unpublished independent project, University of Mysore, Mysore, India.

Sininger, Y., & Starr, A. (2001). Auditory neruropathy: A new perspective on hearing disorder. Canada: Singular publishing group.

Sininger, Y., & Oba, S. (2001). Patients with auditory neuropathy: Who are they and what can they hear? In Y. Sininger, & A. Starr (Eds.), Auditory neuropathy: A new perspective on hearing disorder (pp. 15-36). Canada: Singular publishing group.

Snell, K.B., Mapes, P.M., Hickman, E.D., & Frisina, D.R. (2002). Word recognition in competing babble and the effects of age, temporal processing, and absolute sensitivity. Journal of the Acoustical Society of America, 112, 720-727.

Starr, A., Picton, T.W., & Kim, R. (2001). Pathophysiology of auditory neuropathy. In: Y. Swinger, & A. Starr (Eds.), Auditory neuropthy: A new perspective on hearing disorder (pp. 67-82). Canada: Singular publishing group.

Starr, A., Picton, T.W., Sininger, Y., Hood, L., & Berlin, C.I. (1996). Auditory neuropathy. Brain, 119, 741-753.

Starr, A., McPherson, D., Patterson, J., Don, M., Luxford, W., Shannon, R, Sininger, Y., Tonakawa, L., et al. (1991). Absence of both auditory evoked potential and auditory percepts dependent on timing cues. Brain, 114, 1157-1180

Starr, A., Michalewski, H.J., Zeng, F.G., Brooks, S.F., Linthicum, F. Kim, C.S., Winnier, D., & Keats, B. (2003). Pathology and physiology of auditory neuropathy with a novel mutation in the M P Z gene. Brain, 126, 1604-1619.

Starr, A., Sininger, Y.S., & Praat (2000). Varieties of Auditory neuropathy. Journal of Basic Clinical Physiology and Pharmocology, 11, 215-229.

Starr, S., Sininger, Y.S., Winter, M., Derebery, M.J., Oba, H., & Michalewski, H.J. (1998). Transient deafness due to temperature sensitive auditory neuropathy. Earand Hearing. 7P, 169-179.

Stein, L., Tremblay, K., Pasternak, J., Banerjee, S., Lindeman, K., & Kraus, N. (1996). Brain stem abnormalities in neonates with normal otoacoustic emissions. Seminars in hearing, 17, 197-213.

Strouse, A., Ashmead, D.H., Ohde, R.N., & Gramtham, D.W. (1998). Temporal processing in the aging auditory system. Journal of the Acoustical Society of America, 104, 2385-2399.

Sundberg, U & Lacerda, L. (2003). Does the training with manipulated stimuli improve auditory perception in non typically language leaning children? Phonum, 9, 13-16.

Tallal, P., & Merzenich, M. (1997). Fast Forward training for children with language learning problems: National field test results. Paper presented at the annual meting of- the American Speech Language and Hearing Association, Boston, MA.

Tallal, P., & Piercy. (1975). Developmental aphasia: the perception of brief vowels and extended stop consonants. Neuropsychologia, 13, 69-74.

Tallal, P., Miller, S.T., Bedi, G., Byma, G., Wang, X., Nagarajan, S., Schreiner, C, Jenkins, W., et al. (1996). Language comprehension in language learning impaired children improved with acoustically modified speech. Science, 277, 81-84.

Tang, T.P., Mcpherson, B., Yuen, K.C., Wong, L.L., & Lee, J.S. (2004). Auditory neuropathy/auditory dys-synchrony in school children with hearing loss: Frequency of occurrence. International Journal of Pediatric Otolaryngology, 168, 175-183.

Taylor, M.M., & Creelman, CD. (1967). PEST: Efficient estimate on probability functions. Journal of the Acoustical Society of America, 41, 782-787.

Temple, IS., Deutsch, G.K., Poldrac, R.A., Miller, S.L., Tallal, P., Merzenich, M.M., et al. (2003). Neural deficits in children with dyslexia ameliorated by behavioral remediation: Evidence from functional MRI. Proceedings of National Academy of sciences of the United States of America, USA, 100, 2860-2865.

Trautwein,P.G., Sininger, Y.S., & Nelson, R. (2000). Cochlear implantation of auditory nueropathy. Journal of Anerican Academy of Audiology, 11, 309-315.

Turner, T.W., Souza, P.E., & Forget, L.N. (1995). Use of temporal envelope cues in speech recognition by normal and hearing-impaired listeners. Jouranl of the Acoustical Society of America, 97, 2568-2576.

Tyler, R.S., Summerfield, Q., Wood, E.J., & Famandes, M.A. (1982). Psychoacoustic and phonetic temporal processing in normal and hearing impaired listeners. Journal of the Acoustical Society of America, 72, 740-752.

Uchansky, R.M., Choi, S., Braida, L.D., Reed, CM., & Durlach, N.I. (1996). Speaking clearly for the hard of hearing IV: Further studies of role of speaking rate. Journal of Speech and Hearing Research, 39, 494-509,

Vanaja, C.S., & Jayaram, M. (2003). Sensitivity and specificity of audiological tests in differential diagnosis of auditory disorders. Unpublished project, All India Institute of Speech and Hearing, Mysore, India.

Vandana. (1998). Speech identification test in Kannada. Unpublished independent project, University of Mysore, Mysore, India.

Wang, M.D., & Bilger, R.C. (1973). Consonsant confusions in noise: Astudy of perceptual features. Journal of the Acoustical Society of America, 54, 1248-1266.

Winter, I.M., & Palmer, A.R. (1990). Responses of single units in the anteroventral cochlear nucleus of guinea pig. Hearing Research, 33, 175-180.

Withnell, R, H. (2001). The cochlear microphonic as an indication of outer hair cell function [Brief report]. Ear and Hearing, 22, 75-71.

Zeng, F.G., Kong, Y.Y., Michalewski, H.J., & Starr, A. (2005). Perceptual consequences of disrupted auditory nerve activity. Journal of Neurophysiology, 93, 3050-3063.

Zeng, F.G., Oba, S., Garde, S., Sininger, Y., & Starr, A. (1999). Temporal and speech processing deficits in auditory neuropathy. Neuro Report, 10, 3429-3435.

Zeng, F.G., Oba, S., Garde, S., Sininger, Y., & Starr, A. (2001). Psychoacoustics and speech perception in auditory neuropathy. In: Y. Sininger, & A. Starr (Eds.), Auditory neiropathy: A new perspective on hearing disorder (pp. 141-164). Canada: Singular publishing group.

Zhou, R., Abbas, P.J., & Assouline, J.G. (1995). Electrically evoked auditory brainstem response in peripherally myelin deficient mice, Hearing Research, 88, 98-106.

International Journal of Audiology

Voice 214 905-3001
Fax 214 905-3022

1966 Inwood Road
Dallas, TX 75235

February 6, 2006

MS: "Prevalence and auditory...auditory dys-ynchrony"
MS#: 1009705

Dear Mr. Kumar and Dr. M. Jayaram:

The editorial review on the revised version of the above listed manuscript has been finalized and the paper is accepted for publication. This letter is to inform you that the paper is being sent to the publisher for final production. At present, the production process requires about 5-6 months.

Once again, thank you for submitting this interesting paper. Please feel free to contact the IJA editorial office at editor-ija(%utdallas.edu should you have any questions.

Sincerely,

Ross J. Roeser, Ph.D.
Editor-in-Chief

Prevalence and Audiological Characteristics of Individuals with Auditory Dys-Synchrony

Authors: Ajith Kumar U1 and Jayaram M2

1. JRJF, Department of Audiology, All India Institute of Speech and Hearing, Mysore.
2. Director, All India Institute of Speech and Hearing, Manasagangothri, Mysore.

Corresponding Author:
Mr. Ajith Kumar U
JRF, Department of Audiology,
All India Institute of Speech and Hearing, Naimisham Campus,
Manasagangothri,
Mysore, 570006
INDIA.
Ph: +91 821 2514449 Ext 186 (O)
+91 821 2514449 Ext 225 (R)
Fax:+91 8212510515
Email: ajithkumarl8(5)Email.com

Key words: Auditory dys-synchrony, incidence, otoacoustic emission, and Speech identification scores.

Prevalence and Audiological Characteristics of Individuals with Auditory Dys-synchrony

ABSTRACT

The purpose of this study was to a) estimate the prevalence of auditory dys-synchrony in Mysore, a city of one million population in Southern India, and b) to present the audiological findings, as well as the relationship between different findings, in this clinical population. A register based study design was employed wherein the results of audiological findings of 21,236 patients who visited the Department of Audiology, All India Institute of Speech and Hearing between January 2000 and December 2003 were reviewed. Results showed that the prevalence of auditory dys-synchrony was around 1 in 183 in individuals with sensorineural hearing loss. Behavioral thresholds and speech identification scores were variable. About 60% of the auditory dys-synchronics had no measurable speech identification scores. There was no relationship between hearing thresholds and speech identification scores, or between otoacoustic emissions and speech identification scores in this clinical population. Overall, the results of the present study indicated that auditory dys-synchrony is not an extremely rare disorder.

The use of auditory evoked potentials such as auditory brainstem responses (ABRs) in the assessment of hearing is well established now. A number of studies have demonstrated a strong correlation between ABR thresholds and behavioral hearing levels in both normally hearing and hearing-impaired subjects (Gorga, et al, 1985; Boettcher, 2002; Anthony, et al, 2002, among others). However, there have been reports in literature on subjects in whom evoked potential thresholds are significantly worse than can be expected from their audiogram (Davis & Hirsh, 1979; Kraus, et al, 1984; Starr et al, 1996). These patients may not benefit from hearing aids. Starr et al (1996) used presence of preneural responses such as cochlear microphonics and otoacoustic emissions in a group of 10 individuals with absent ABR to identify normal functioning outer hair cells. The term 'auditory neuropathy' or 'auditory dys-synchrony' is used to denote the anatomic and functional disturbances in these individuals. Advances in the accurate measurement of outer hair cell function have made it easier to diagnose this hearing disorder characterized by abnormal temporal encoding (Zeng, et al, 1999) and neural asynchrony.

Davis and Hirsh (1979) reported that 1 in 200 hearing impaired children exhibit an audiological picture that is consistent with the contemporary diagnosis of auditory dys-synchrony. The overall prevalence rate varies from 1 1% (Ranee, et al, 1999 - hospital based statistics) to 1.83% of hearing impaired population (Tang, et al, 2004 - on school-aged population). Both these studies are on Western population and have used a relatively small sample size (Ranee, et al, 1999 - N300; Tang, et al, 2004 - N56).

The exact site of lesion and pathophysiology of auditory dys-synchrony is not yet completely understood. These individuals have preserved cochlear amplification, but disturbed normal synchronous activity of the auditory nerve.

Discharges of the auditory nerve are presumed to be asynchronous as is evident from normal otoacoustic emissions in the presence of absent or abnormal auditory brainstem responses. These individuals typically have speech recognition deficits that are not in consonance with their puretone hearing thresholds. They do not usually benefit from conventional amplification. Poor speech perception in these patients is attributed to abnormal temporal coding and asynchrony (Zeng et al, 1999; Kraus et al, 2000; Ranee et al, 2004; Zeng et al, 2004). Etiologies of auditory dys-synchrony are only now beginning to be appreciated and appear diverse. It has often appeared in clinical reports that neonates, at risk for hyperbilirubinemia and anoxia, seem to be at risk for auditory dys-synchrony as well (Ranee et al 1999). Genetic factors have also been identified. Starr et al (2003) reported a novel mutation in M P Z gene in a family with hereditary motor sensory neuropathy and deafness.

Intervention strategies for individuals with auditory dys-synchrony are very different from those provided to many other hearing impaired children. Conventional amplification through hearing aids does not seem to be beneficial as this does not address the problem of neural dyssynchrony. Usefulness of hearing aids is also reported to be dependent on cortical evoked potentials and temporal processing abilities (Ranee et al., 2002, 2004). Cochlear implantation is of benefit to some patients with auditory dys-synchrony (Sininger, et al, 1999). However, the usefulness of cochlear implantation seems to depend on the site of lesion. Communication difficulties in individuals with auditory dys-synchrony, even in those with mild hearing loss, are much more severe compared those individuals with cochlear hearing loss of 60 dB HL or more. Even then, individuals with auditory dys-synchrony are not covered under the disabled category [for example: In India, Persons with Disability Act (1995)].

Therefore, these individuals cannot avail the benefits which are otherwise provided for individuals with hearing loss of 60 dB HL or more.

It is necessary to know the prevalence, audiological characteristics and the relationship between different audiological measures in this clinical population to reformulate the rules governing their hearing problem and to select appropriate management strategies. Therefore, the present study aimed to

1) estimate the prevalence of auditory dys-synchrony in hearing impaired individuals with sensori-neural hearing loss, and

2) to analyze the relationship between speech identification scores and other audiological measures such as puretone threshold, amplitude of otoacoustic emission, and audiometric configuration in this group of patients with auditory dys-synchrony..

METHOD

Subjects

A register based study design was employed to find the prevalence and to describe the audiological characteristics of patients with auditory dys-synchrony. A retrospective analysis of case reports was done wherein test findings of all those who visited the Department of Audiology, All India Institute of Speech and Hearing, Mysore between January 2000 and December 2003 were reviewed. Mysore is city in Southern India and has a population predominantly rural of about a million. People in this part of the country speak Kannada a Dravidian language (Jayaram, 1984) A total of 21,236 (11,712 males and 9524 females) records were reviewed. 11,205 of these were of persons with permanent sensori-neural hearing loss. 61 of these 11,205 (5854 males and 5351 females) individuals were identified as cases of auditory dys-

synchrony. Criteria employed to identify auditory dys-synchrony were those recommended by Starr, et al (2000). Accordingly patients who showed

a) preserved cochlear amplification, that is, presence of transient evoked otoacoustic emissions,

b) altered auditory nerve responses (absent or severely abnormal auditory brainstem responses),

c) no evidence of space occupying lesion on neurological examination (clinical neurological examination or CT scan or MRI) in the presence of normal otological and tympanometric findings were identified to have auditory dys-synchrony.

All these individuals had undergone a) puretone, speech and immittance audiometry, and b) auditory brainstem response and otoacoustic emissions evaluation. However, otoacoustic emission testing had been done only when ABR was abnormal. Roughly, 20% of the patients had undergone neurological examination, but all the 61 individuals who were diagnosed to have auditory dys-synchrony had been subjected to all tests including neurological investigation.

Testing procedure and instruments

It was ascertained from case records that all these subjects had been tested under standard conditions. All subjects had been tested with calibrated (ISO, 983) audiometers in sound treated rooms. Puretone testing had been done using modified version of Hughson and Westlake procedure. Speech identification testing had been done with live voice presentation of phonetically balanced monosyllables at maximum comfortable level. Immittance evaluation (tympanometry and acoustic reflex threshold testing) for 226 Hz probe tone had been carried out with calibrated middle ear analysers (GSI-33 or Tympstar).

Auditory brainstem response measurement

Auditory brainstem response testing had been done using either Biologic Navigator or Nicolet Bravo evoked potential system. Identical protocol had been used to test all the patients. Table 1 shows the protocol used for auditory brainstem response testing. It was checked from the records that auditory brainstem testing had been done twice to ensure reproducibility of waveforms. A group of 30 normal hearing adults had been tested to establish the nHL values. Results showed that 0 dB nHL click had a peak equivalent SPL of 26 dB. Transient evoked otoacoustic emissions (ILO 292) were measured in a sound treated room for clicks at 80 dB + 5 dB pe SPL. An emission was considered to be present if the waveform reproducibility was more than 50% and the overall signal to noise ratio was more than 3 dB at two frequency bands at least.

Neurological evaluation

All the 61 patients had been subjected to a neurological examination by a qualified neurologist. This included clinical neurological examination, and/or CT scan and/or MRI.

RESULTS

Prevalence

61 of 21,236 hearing impaired were identified to have auditory dys-synchrony following the criteria of Starr et al (2000) which means that 1 out of 348 hearing impaired had auditory dys-synchrony. However, the prevalence was 1 in 183 (0.54%) when only individuals with permanent sensoh-neural hearing loss were considered (61 of 11,205).

The average age of onset was 16 years (range 1 to 31 years). 59% of the 61 patients with auditory dys-synchrony had onset of the problem between the

age of 14 to 24 years (Figure 1). 81% of this auditory dys-synchrony population reported their hearing problem as progressive. No specific etiology could be traced in most of these patients. Two patients had a neurodegenerative disorder and 2 patients reported that the problem started after they gave birth to their first child.

Insert Figure 1

Only two patients with auditory dys-synchrony demonstrated abnormal MR1, while the remaining presented normal clinical neurological and imaging studies. All patients showed normal findings on otologic examination. There was no evidence of any middle ear disorder. 50% of the auditory dys-synchrony population reported bilateral tinnitus while 8 patients reported of vertigo. All patients who had vertigo complained of tinnitus also.

Puretone audiometry

In all the 61 patients, the auditory problem was bilateral though asymmetry was observed in some individuals. Females and males were affected in the ratio of 2:1. 54 of the 61 the patients, came from poor socioeconomic strata (monthly income Rs < 2000 or < USD 42). Of the 61 patients, 26 showed peaked audiograms (sharp peak at a single frequency with worsening of thresholds at immediately adjacent frequencies), 11 showed flat (<5 dB raise or fall per octave), 11 showed rising (a 5 dB or more decrease in threshold per octave), 8 showed saucer-shaped audiograms (20 dB or more loss at the extreme frequencies compared to middle frequencies) and 3 showed sloping audiograms (5-12 dB increase in threshold per octave) in both ears (Figure 2). Among patients who showed peaked audiogram, 77% had a peak at 2000 Hz. Puretone thresholds were not available in two children below 2 years of age.

Insert Figure 2 here

Degree of hearing loss varied from mild to severe. This classification may not be valid in most of the cases because not only were the responses inconsistent, but also a majority of them had peaked audiograms. Responses of 40% of the patients were judged as inconsistent because thresholds had varied by more than 10 dB within a test session. Paired two-tailed T test showed no significant difference between the two ears with respect to 3-frequency pure tone average ($t = 0.5$, $p > 0.05$).

Speech perception

Speech perception abilities of patients varied from no measurable speech identification score to 90% speech identification score. Figures 3a and 3b show the distribution pattern of speech identification scores for the right and the left ear, respectively. It is clear from Figures 3a and b that around 60% of the patients did not have measurable speech identification scores.

Insert Figures 3a and 3b here.

Relation between the speech identification scores and the results of other auditory assessments were examined in all the patients. Patients' thresholds at each of the audiometric frequencies (octave frequencies between 250 Hz and 8 kHz) were compared with speech perception score by computing a Pearson's Product Moment Correlation between the two factors. There was a significant negative correlation between the two variables at all the frequencies except at 8 kHz. The observed correlation may have been because of subjects who had no measurable speech perception abilities in the presence of mild to moderate

Insert Figures 4 here.

hearing loss. A scatter plot was drawn between the two variables, at all frequencies, to verify the validity of correlation (Figure 4).

The correlation coefficient 'r' is shown on the top right comer of each of the graphs. An analysis of the scatter plots revealed no specific trend between hearing threshold, at any frequency, and speech identification scores except at 250 Hz and 500 Hz (individuals who had no measurable speech identification scores were excluded form the scatter plot). Speech perception scores were also compared with audiogram configuration. Figure 5 shows the mean speech identification scores for each of the audiometric configurations described earlier. It is clear from this figure that patients who had peaked audiogram showed higher speech identification scores, in both ears, than patients with other audiometric configurations. But, any generalization on this is dangerous because of the far too small number of patients, particularly those with rising and saucer shaped audiograms.

. Immittance evaluation and auditory brainstem responses

Immittance evaluation showed normal tympanograms, but absent stapedial acoustic reflexes in all patients. None of the patients showed A B R for clicks in either ear.

Otoacoustic emissions

Transient evoked otoacoustic emissions were bilaterally present in all the 61 subjects. The mean emission amplitude of T E O A E was 16.7 dB SPL (SD = 3.9) for the right ear and 16.4 dB SPL (SD = 3.6) for the left ear. A two-tailed paired T test failed to show any statistically significant difference between the right and the left ear in the amplitude of otoacoustic emissions ($t = 1.23$, $p > .05$). The relation between T E O A E amplitude and puretone thresholds at 500 Hz, 1 kHz, 2 kHz and 4 kHz was examined. As there was no statistically significant

difference between the two ears in terms of TEOAE amplitude and puretone threshold, data from both the ears were combined. Pearson's Product Moment Correlation was not significant between hearing threshold and TEOAE amplitude at any frequency. Correlation between the amplitude of TEOAEs and speech identification scores was also not significant for any ear.

DISCUSSION

61 of 11, 205 individuals with permanent sensory-neural hearing loss were identified to have auditory dys-synchrony, which means a prevalence of 0.54% or 1 in 183 persons. This figure is significantly lower than the figure of 11% reported by Ranee et al (1999) and the 1.8% reported by Tang et al (2004). This difference might be because of the difference in the population tested. A large number of subjects in the study of Ranee et al (1999) were bom premature and were at risk for neurodevelopmental disorder (hyperbilirubinemia). It is known that hyperbilirubinemia can cause both permanent and temporary dysfunction of auditory pathways. Some of the infants may recover from abnormal ABR as they grow. Therefore, the figure reported by Ranee et al (1999) seems to be an overestimate of the extent of auditory dys-synchrony. The figure of 1.8% reported by Tang et al (2004), in school aged children (N=56), also seems to be an over estimation because sample size is too small to allow any generalization. In the 56 children that they studied, only one child had evidence of auditory dys-synchrony which comes to 1.8%. The results of the present study are nearer to those reported (1/200) by Davis and Hirsh (1979). Actually, the prevalence of auditory dys-synchrony may be slightly higher than that is reported in the present study because some or the other test result was not available for nearly 30% of the hearing impaired in the present study, and thus, some subjects might have missed the diagnostic criteria for auditory dys-

synchrony employed in this study. Also, the present study employed transient evoked otoacoustic emissions to infer normal functioning of outer hair cells. However, Starr et al (2000) have reported that around 16% of individuals with auditory dys-synchrony would loose otoacoustic emissions over time. Thus, some patients with auditory dys-synchrony might have been missed out in the present study. However, irrespective of the number of patients with auditory dys-synchrony, early identification of this disorder is necessary because conventional management technique like use of amplification devices does not seem to be effective with this population and sometimes may even be dangerous.

Only 2 of the 61 patients in the present study showed neurological deficits which is much lower than the 80% reported by Starr et al (2000). This may be because Starr et al (2000) dealt predominately with neurological population in a medical setup. However, the All India Institute of Speech and Hearing, Mysore where the present study was conducted functions in a non-medical setup. Thus many patients with medical history may not have consulted at this institute.

Female to male ratio of auditory dys-synchrony was 2:1 in this study. This along with the report of two patients that they developed the problem after the delivery of their first child suggests the need for investigation of hormonal influence in precipitating the condition of auditory dys-synchrony.

The most common audiometric configuration seen in persons with auditory dys-synchrony in this study was 'peaked' type, with the peak usually occurring at 2 kHz. Frequent occurrence of such an audiogram may be due to anatomico-physiological make up of the auditory nerve. The longest fibers are most susceptible to pathology. The longest cochlear nerve fibers are those going

cochlea/auditory nerve fibers at mid frequencies might be responsible for poor speech identification scores, in addition to temporal asynchrony in auditory nerve firings, in patients with rising, sloping and saucer-shaped audiograms than in patients with peaked audiograms. Even though there was a significant negative correlation between thresholds at all frequencies and speech identification scores except at 8 kHz, scatter plots showed no specific trend between the two variables except at low frequencies. No relation between speech identification scores and behavioral thresholds are expected either as inability of theses patients to understand speech is because of temporal asynchrony in auditory system and not due to elevated thresholds, per se.

This frequency specific correlation may be related to the underlying physiology of high frequency and low frequency coding. Low frequencies are usually coded by phase locked responses in type I auditory nerve fibers. Individuals with auditory dys-synchrony cannot use phase locking cues to the same extent as normal hearing listeners due to dyssynchronous firing of auditory nerve fibers. However, detection ofhigh frequencies depends on information on place of excitation on basilar membrane, and does not depend on phase locking cues as much as low frequencies. We hypothesise that low frequency hearing sensitivity in these individuals may indicate the extent of temporal disruption in the auditory system. Hence, the greater the loss in low frequencies, the greater is the severity of temporal asynchrony which in turn reduces speech perception abilities.

Mean amplitude of TEOAE for adults in our clinic is 11.5 dB SPL. The amplitude of TEOAE in individuals with auditory dys-synchrony in this study was 16 dB SPL. Higher amplitude of evoked otoacoustic emission in patients with auditory dys-synchrony has also been reported by others (Hood & Berlin,

2001). This phenomenon has been attributed to lack of efferent suppression of otoacoustic emissions. As there is evidence to say that the efferent system helps in the perception of speech (Muchink, et al, 2004), and high amplitudes of TEOAE reflecting the non-functioning of efferent system, one would expect poor speech identification scores to be correlated with high amplitudes of TEOAE. However, this was not observed in the present study. Amplitude of TEOAE also did not correlate with behavioral threshold at any of the frequencies. Hence, OAEs are not helpful in predicting speech identification scores or behavioral thresholds in patients with auditory dys-synchrony.

In summary, the results showed that the prevalence of auditory dys-synchrony in individuals with permanent sensori-neural hearing loss is 1 in 183. More females than males are affected. The subjects described in this study varied significantly in their audiological characteristics. The most common configuration seen in these patients was a 'peaked' audiogram. Speech identification scores varied widely. Generally individuals with 'peaked' audiograms had better speech identification scores. There was no correlation between speech identification scores and other audiological measures tested.

REFERENCES

Anthony, T. C, Joaquim, M.B. & Pinherio. 2002. Relation between otoacoustic emission and auditory brainstem responses in neonates and young children: A correlation and factor analytical study. .Laryngoscope, 112,156-167.

Boettcher, F.A. 2002. Presbyacusis and auditory brainstem responses. J Speech Lang Hear Res, 45, 1249-1261.

Davis, H. & Hirsh, S.K. 1979. A slow brainstem response for low frequency audiometry.Audiology, 18,445-465.

Halpin, C. 2003. Tuning curve audiograms. Am J Audiol, 11, 56-64.

Hood, L.J. & Berlin, C.I. 2001. Auditory neuropathy/ (auditory dys-synchrony) disables efferent suppression of otoacoustic emissions. In: Y.Sininger & A. Starr (eds.) Auditory neuropathy: A new Prespective on hearing disorder Canada: Singular publishing group, pp. 183-202.

Gorga, M.P., Worthington, D.W., Reiland, J.K., Beauchaine, K.A. & Goldgar, D.E . 1985. Some comparisons between the auditory brain stem response threshold, latencies, and the puretone audiogram. Ear .Hear, 6,105-112.

ISO 3 89-1:1998 Reference zero for the calibration of audiometric equipment - Part 1: Reference equivalent threshold sound pressure levels for puretones and supra-aural earphones.

Jayaram, M. 1984. Distribution of stuttering in sentences: Relationship to sentence length and clause position. J Speech Hear Res, 27, 38-42.

Kraus, N., Ozdamar, O., Stein, L. & Reed, N. 1984. Absent auditory brainstem response: Peripheral hearing loss or brainstem dysfunction. L:aryngoscope, 94, 400-406.

Kraus, N., Bradlow, A.R., Cheatham, M.A., Cunningham, J., King, CD., Koch, D.B., Nicol, T.G., McGee, T.J., Stein, L.K. & Wright, B.A. 2000.

Consequences of neural asynchrony: A case of auditory neuropathy. J Assoc Res Otolaryngol, 1, 33-45.

Muchink, C, Roth, D.A.V., Jebara, R.O., Katz, H.P., Shabtai, E.L. & Hildesheimer, M. 2004. Reduced medial olivocochlear bundle system function in children with auditory processing disorders. Audiol Neurootol, 9, 107-114.

Rance, G., Beer, D.E., Cone-Wesson, B., Shepherd, R.K., Dowell, R.C. & King, A.M. 1999. Clinical findings for a group of infants and young children with auditory neuropathy. Ear Hear, 20, 238-252.

Rance, G., Cone-Wesson, B., Wunderlich, J. & Dowell, R. 2002. Speech perception and cortical event related potentials in children with auditory neuropathy. Ear Hear. 23, 239-253

Rance, G., McKay, C. & Grayden, D. 2004. Perceptual characterization of children with auditory neuropathy, Ear Hear. 25, 34-46.

Sininger, Y., Trautwein, P., Shallop, J., Fabry, L. & Starr, A. 1999. Electrical activation of auditory nerve in patients with auditory neuropathy. Assoc Res Otolaryngol abstracts, 22,170.

Starr, A., Michalewski, H.J., Zeng, F.G., Brooks, S.F., Linthicum, F., Kim, C.S., Winnier, D. & Keats, B. 2003. Pathology and physiology of auditory neuropathy with a novel mutation in the M P Z gene. Brain, 126, 1604-1619.

Starr, A., Picton, T.W. & Kim, R. 2001. Pathophysiology of auditory neuropathy. In: Y. Sininger & A. Starr (eds.) Auditory neuropathy: A new Perspective on hearing disorder Canada: Singular publishing group, pp. 67-82.

Starr, A, Sininger, Y.S. & Praat, H. 2000. The varieties of auditory neuropthy. J Basic Clin Physiol Pharmacol, 11, 215-229.

Starr, A., Picton, T.W., Sininger, Y., Hood, L.J. & Berlin, C.I. 1996. Auditory neuropathy. Brain, 119, 741-753.

Starr, A., Sininger, Y., Winter, M., Derebery, M.J., Oba, H. & Michalewski, H.J. 1998. Transient deafness due to temperature sensitive auditory neuropathy. Ear hear. 19, 169-179.

Tang, T.P.Y., McPherson, B., Yuen, K.C.P., Wong, L.L.N. & Lee, J.S.M. 2004. Auditory neuropathy/Auditory dys synchrony in school children with hearing loss: frequency of occurrence. Int J Pediatr Otorhinolaryngol, 68,175-183.

Zeng, F.G., Oba, S., Sininger, Y. & Starr, A. 1999 i Temporal and speech - processing deficits in auditory neuropathy. Neuroreport, 10, 3429-3435.

Zeng, F.G., Kong, Y.Y., Michalewski, H.J. & Starr, A. 2004. Perceptual consequences of disrupted auditory nerve activity. J Neurophysiol, . 93, 3050-3063

Behavioral and Brain Functions Bio Med Central

Research Open Access

Auditory processing in individuals with auditory neuropathy

Ajith U Kumar* and M Jayaram[2]

Address: junior Research Fellow, Department of Audiology, All India Institute of Speech and Hearing, Manasagangothri, Mysore, Karnataka, 570006, India and [2]Director, All India Institute of Speech and Hearing, Manasagangothri, Mysore, Karnataka, 570006, India

Email: Ajith U Kumar* - ajithkumarl8@gmail.com; M Jayaram - aiish_dir@yahoo.com

* Corresponding author

Published: 01 December 2005 Received: 06 June 2005

Behavioral and Brain Functions 2005, 1:21 doi: 10.1186/1744-9081 -1 -21 Accepted: 01 December 2005

This article is available from: http://www.behavioralandbrainfunctions.eom/content/l/l/2l

Abstract

Background: Auditory neuropathy is a disorder characterized by no or severely impaired auditory brainstem responses in presence of normal otoacoustic emissions and/or cochlear microphonics. Speech perception abilities in these individuals are disproportionate to their hearing sensitivity and reported to be dependent on cortical evoked potentials and temporal processing abilities. The disproportionate loss of auditory percept in presence of normal cochlear function is suggestive of impairment of auditory neural synchrony.

Methods: We studied the auditory evoked potentials and psychophysical abilities in 14 adults with auditory neuropathy to characterize their perceptual capabilities. Psychophysical tests included measurement of open set speech identification scores, just noticeable difference for transition duration of syllable /da/ and temporal modulation transfer function. Auditory evoked potentials measures were, recording of $P|/N_h$ P_2/N_2 complex and mismatch negativity (MMN).

Results: Results revealed a significant correlation between temporal processing deficits and speech perception abilities. In majority of individuals with auditory neuropathy $P|/N_h$ P_2/N_2 complex and mismatch negativity could be elicited with normal amplitude and latency. None of the measured evoked potential parameters correlated with the speech perception scores. Many of the subjects with auditory neuropathy showed normal MMN even though they could not discriminate the stimulus contrast behaviorally.

Conclusion: Conclusions drawn from the study are

1. Individuals with auditory neuropathy have severely affected temporal processing.

2. The presence of MMN may not be directly linked to presence of behavioral discrimination and to speech perception capabilities at least in adults with auditory neuropathy.

Background

Auditory neuropathy (AN) is recently described hearing disorder characterized by abnormal auditory nerve functioning in presence of normal cochlear receptor hair cell activity [1]. The clinical findings that define auditory neuropathy are

Table 1: Audiometric and electrophysiological details of auditor neuropathy subjects.

SN	Age/sex	PTA (.5,1 and 2 KHz)	Speech identification scores	OAE	ABR	Acoustic reflex	Efferent suppression	Configuration
1	16/F	45.00	45.00	Present	Absent	Absent	0.2	Raising
2	16/M	13.00	84.00	Present	absent	Absent	0.0	Peaked
3	30/M	23.00	80.00	Present	absent	Absent	0.4	Peaked
4	24/F	25.00	38.00	Present	absent	Absent	0.0	Peaked
5	16/M	40.00	.00	Present	absent	Absent	0.1	Raising
6	26/M	45.00	4.00	Present	absent	Absent	0.0	Raising
7	23/F	45.00	5.00	Present	absent	Absent	0.0	Raising
8	23/M	75.00	.00	Present	absent	Absent	0.1	Flat
9	27/M	20.00	50.00	Present	absent	Absent	0.0	Peaked
10	23/M	40.00	86.00	Present	absent	Absent	0.3	Peaked
11	24/M	23.00	8.00	Present	absent	Absent	0.0	Peaked
12	25/F	45.00	.00	Present	absent	Absent	0.0	Raising
13	28/M	5.00	90.00	Present	absent	Absent	0.3	Peaked
14	25/F	10.00	95.00	Present	absent	Absent	0.2	Peaked

PTA = Pure tone average
OAE = Otoacoustic emissions
ABR = Auditory brainstem responses

a) Presence of outer hair cell integrity in evoked otoacoustic emission or cochlear microphonics.

b) Absence of synchronized neural activity at the level of 8th nerve and brainstem.

Though the audiometric and electrophysiological findings are consistent with the 'retro outer hair cell dysfunction' exact site(s) of the pathology is yet to be determined. Some possible sites of lesion that could produce the audiometric and electrophysiological profile of AN include: inner hair cells, synaptic junction between inner hair cell and type I afferent nerve fibers, spiral ganglion cells, specific damage or demyelinization of type I auditory nerve fibers [1-3]. Therefore, AN consists of many varieties depending on the sites of lesion [4]. Speech perception ability in these patients also varies considerably. Some patients perform at the levels expected for patients with comparable degrees of sensory hearing loss and others show speech understanding which is disproportionate to their degree of hearing loss [5,6].

Speech perception abilities in these patients appear to depend on the extent of suprathreshold temporal distortions of cues rather than access to speech spectrum, unlike the patients with sensory hearing loss [7,6]. Zeng et al [8] reported the abnormal results on two measures of temporal perception in their group of children with AN: (i) gap detection threshold (identification of silence embedded in within the bursts of noise) and (ii) temporal modulation transfer function (measure of sensitivity to slow and fast amplitude fluctuation). They also found a correlation between temporal modulation transfer function (TMTF) and speech perception abilities in their patients. Ranee et al [6] also reported poor performance on the task involving timing cues (TMTF, temporal aspects of frequency discrimination) in a group of 14 children with AN. These temporal processing abnormalities had significant correlation with speech perception abilities. They attributed the speech perception scores that are disproportionate to pure tone hearing loss to these suprathreshold temporal processing deficits.

Another factor that is reported to be related to speech perception abilities in these individuals is cortical evoked event related potentials. Ranee et al [5] reported that a subgroup of children with AN, who had recordable cortical evoked potential performed well on open set speech perception task and derived significant benefit from amplification. In contrast, subjects who had no recorda ble cortical evoked potential performed poorly on the same tasks. From this observation they concluded that presence of cortical auditory evoked potential reflects some amount of preserved synchrony in central auditory system which contributes to better speech understanding despite the distortion that occurs at 8th nerve and auditory brainstem in these individuals.

Speech perception process can be investigated in neurophysiological as well as psychophysical perspective. An important aspect of this study is use of a combined neurophysiological and psychophysical approach. With this multidisciplinary technique we hope to gain insight into both stimulus representation and processing in individuals with AN. This study is sought to explore the relation between their psychoacoustic abilities and evoked poten-

tial parameters, in a group of adults with auditory neuropathy. Psychophysical experiments included were measurement of open set speech identification scores, just noticeable difference (JND) for transition duration of the syllable /da/ and temporal modulation transfer function. Auditory evoked potentials measures included recording of N_1/P_1, N_2/P_2 and Mismatch negativity (MMN) potentials.

Methods

Study was carried out in two phases, first phase involved psychophysical experiments and auditory evoked potentials were measured in the second phase.

Subjects

Two groups of subjects participated in the study. The first group consisted of 14 individuals with AN (16 to 30 years with the mean age of 23 years) and second group consisted of age and gender matched 30 normally hearing subjects. All AN subjects were recruited from Department of Audiology, All India Institute of Speech and Hearing, Mysore. No subject complained about any middle ear disease (assessed using otoscopy, tympanometry and clinical history), noise exposure or ototoxic drug usage. Results of different audiological measurements of AN subjects are shown in Table 1. As all the subjects had symmetrical hearing loss, (symmetrical hearing loss was operationally defined as the difference in thresholds between two ears at corresponding frequencies within 15 dB), pure tone thresholds were measured again with loudspeakers and these measurements were considered for all future purpose. Furthermore, subjects in the normally hearing group had their hearing thresholds within 15 dB HL at octave frequencies between 250 Hz to 8 kHz and normal results on immittance evaluation. All the subjects were native speakers of Kannada, a South Indian Dravidian language.

Psychophysical tests

The experiment protocol consisted of speech identification score testing, measurement of JND for transition duration of /da/ and TMTF.

(a) Speech identification testing

Only AN subjects participated in this experiment. Vandana's speech identification test in Kannada was used to assess the open set speech perception abilities in the subjects. This test consists of 50 bisyllabic meaningful words in Kannada. Validity and reliability of this test on native speakers of Kannada have already been established by Vandana, [9]. Recorded material was presented at 'comfortable level' which ranged between 30 to 40 dB SL ref: Average thresholds at 500 Hz, 1 kHz, and 2 kHz, using MA-53 clinical audiometer through a loudspeaker kept at 1 m distance and 0° azimuth. Output of the loudspeaker was calibrated using Quest 1800 sound level meter and Quest 4180 free field microphone. A calibration tone recorded before the test material was used to adjust the Vu meter deflection to zero. The test was carried out in a quiet listening condition and each stimulus was presented in isolation without being embedded in a carrier phrase. The subjects were required to repeat each stimulus and a percentage of correct identification was determined. All the subjects were screened for misarticulations using Kannada Articulation Test [10]

(b) JND measurements

Both AN and normal listeners participated in this experiment. Stimulus was derived from retroflex /da/ uttered in isolation, by a 25 year old male native speaker of Kannada. The spoken was digitally recorded on a data acquisition system at 44 kHz sampling frequency. The transition duration was identified using both spectral and wave form view of the stimulus. Transition duration was lengthened up to 'original transition duration +100 ms' in 10 ms steps by means of Pitch Synchronized Overlap and Add (PSOLA) technique. PSOLA performs the lengthening of the stimulus in time domain and preserves most of physical characteristics of the stimulus such as spectral shape, amplitude distribution, and periodicity [11].

Subjects were tested individually in a sound attenuated room. Signals were played via a PC, at a sampling frequency of 44 kHz and were subsequently fed to a MA-53 audiometer. Subjects received the signals through audiometer's loudspeaker kept at a distance of 1 m and 0° azimuth. Presentation level of the stimulus was fixed at 30 dB SL ref: Average thresholds at 500 Hz, 1 kHz, and 2 kHz. Stimuli were presented at equal presentation level to compensate for the audibility in individuals with auditory neuropathy. JND was determined using an adaptive tracking technique (PEST) with AX same difference discrimination paradigm (in this A = anchor stimulus, X = Variable stimulus and subjects task is to indicate whether A is same as X or not). Inter stimulus interval between anchor and variable stimulus was 500 ms. Step size and the direction of variable stimulus were changed according to rules of PEST [12]. The subject's JND was determined by calculating the difference in transition duration between anchor and variable stimuli that is required to achieve a performance level of 69% correct responses. Test trials also included equal number of catch trials. Catch trial consisted of either two identical anchor or two identical non anchor stimuli.

(c) Temporal modulation transformer function

Both AN and normal listeners participated in this experiment. Modulation detection thresholds were measured by determining the sensitivity to sinusoidal amplitude modulation as a function of modulation frequency. Presentation level of the stimulus was kept at 30 dB SL ref: Average

Table 2: Protocol used for evoked potential testing

Stimulus	Standard - unmodified /da/ Deviant — Synthesized /da/
Intensity	30 to 40 dB SL
Probability	5:1
Repetition rate	1.1 /s
Analysis time	500 ms
Gain	75000
Band Pass Filter	1 to 30 Hz
Transducer	EAR-3A insert ear phones

thresholds at 500 Hz, 1 kHz, and 2 kHz. Stimulus was presented through a loud speaker kept at a distance of 1 m and 0° azimuth. Stimuli were presented at equal sensation level to compensate for the audibility in patients with auditory neuropathy. A broad band noise was generated and controlled digitally to measure TMTF. Broad band noise had a duration of 500 ms and ramp of 2.5 ms. The modulated signal was derived by multiplying the 500 ms white noise by a dc shifted sine wave. The depth of the modulation was controlled by varying the amplitude of modulating sine wave. Modulation depth for the various stimuli varied between 0 to -30 dB and step size was 3 dB. Modulation detection thresholds were measured for 5 frequencies; 4 Hz, 16 Hz, 32 Hz, 64 Hz, 128 Hz, and 200 Hz,. Procedure was same as that described for the measurement of JNDs. In all the subjects at least at one modulation frequency the presentation level was changed and modulation detection threshold was rechecked to ensure that subjects are not using the loudness judgments.

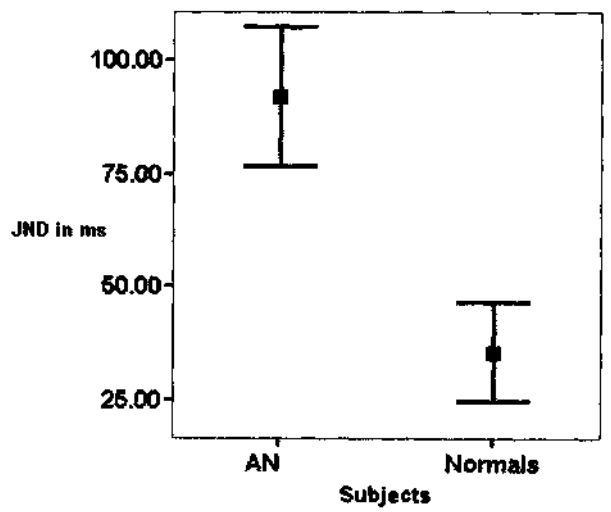


Figure 1
Mean and SD (error bars show 1 SD) of JND in transition duration for the auditory neuropathy (AN) group and normally hearing subjects.

Auditory evoked potential *measurements*

In this experiment both normal hearing subjects and individuals with AN participated. The cortical evoked potentials were obtained in one session lasting less than 15 min. The subjects were seated in a comfortable position to ensure relaxed posture to minimize muscular artifacts. They were instructed not to pay attention to the stimuli. A silent cartoon movie was played to minimize the possibility of active attention. The stimuli was unmodified /da/ and synthesized /da/ in which transition duration was lengthened by 100 ms. This was decided on the basis of a pilot study measuring the behavioral JND in AN subjects. Synthesis technique was same as the one used for psychoacoustic testing. These two contrasts were presented in an odd ball paradigm. Stimuli were presented at 'comfortable level' to both ears (usually 30 to 40 dB SL, Ref: Average thresholds at 500 Hz, 1 kHz, and 2 kHz) through EAR-3A insert receiver. IHS smart EP module was used to control the stimulus presentation and acquisition of evoked potential. Conventional recording techniques were used. After skin preparation at electrode site, silver-chloride disc electrodes were placed at C_z, with ipsilateral mastoid as reference, using conductive electrode paste and adhesive tape. Ground electrode was placed at F_z Data was acquired after ensuring that the impedance at all electrode sites was within permissible limits. The protocol used for recording is shown in Table 2.

In order to probe the representation of these two stimulus contrasts at pre-attentive neural level MMN responses were derived from recorded cortical evoked potentials. MMN is a passively elicited cortical evoked potential that is known to reflect the brain's response to an acoustic change [13]. The MMN is seen as a negative deflection around 200 ms after stimulus presentation. MMN was identified in the difference wave between frequent and infrequent recordings. Grand average waveform was also constructed by utilizing the individual waveform which had MMN.

Results

Psychophysical tests

(a) Open set speech identification test

Open set speech identification scores in individuals with AN varied considerably. The mean speech identification score was 41.7% (SD: 38.8%), but scores ranged from 0% to 95%. Speech identification scores correlated with low frequency (250 Hz, 500 Hz, 1000 Hz) hearing thresholds ($r = 0.67$, $p = 0.001$) but not with the high frequency hearing thresholds (2000 Hz, 4000 Hz and 8000 Hz $r = 0.3$, $p = .234$).

(b) JND measurements

Figure 1 shows the mean and SD values of JND in transition duration for stimuli /da/. Independent sample 't' test

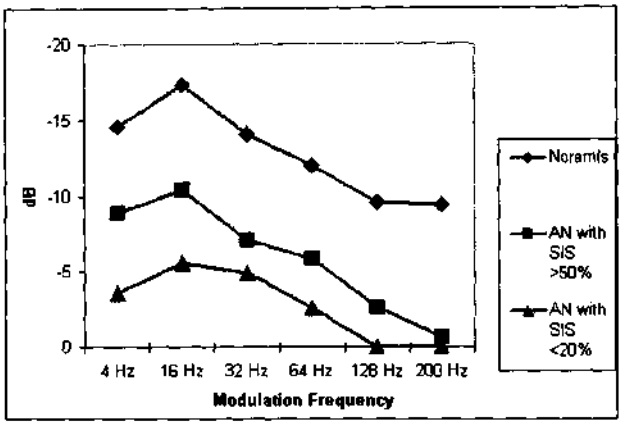


Figure 2
TMTF for the auditory neuropathy (AN) group and normally hearing subjects. AN20 = TMTF for auditory neuropathy subjects with speech identification scores less than 20%. AN50+ = TMTF for auditory neuropathy subjects with speech identification scores more than 50% Normal = TMTF for normally hearing subjects.

showed a significant difference between two groups at .001 level. Of 14 subjects 10 could not differentiate the stimuli that differed in transition duration by as much as 100 ms. Four subjects whose JND were less than 100 ms also had their open set speech identification scores more than 80%.

Temporal modulation transfer function

Figure 2 shows the TMTF for subjects with normal hearing and auditory neuropathy. Normal hearing listeners were most sensitive to slow temporal fluctuation and became less sensitive as the fluctuation rate was increased. Similar trend was noticed in individuals with AN. Average peak sensitivity of normal hearing listeners was -17.36 dB. In contrast, average peak sensitivity for auditory neuropathy group was -6.6 dB (SD: 5.4 dB). At higher modulation frequencies many of the AN (12 subjects) subjects did not even detect a modulation of depth of 0 dB (100%). Peak sensitivity of AN group tended to fall in two distinct categories. Eight individuals had peak sensitivity of more than -10.4 dB and 7 of these patients had open set speech identification scores more than 50%. Six subjects had peak sensitivity less than -5.6 dB and 5 of them had speech identification scores of less than 20%. One subjects i-each category had paradoxical results on speech perception and TMTF results. When data from individual subjects were examined speech identification scores and temporal modulation transfer function in these two subjects were in extreme. Hence these two subjects were treated as outliers and when data from these two subjects were excluded a significant correlation was observed between peak sensitivity and speech identification scores. No relation could be established between JND measurements and TMTF.

Auditory *evoked potential measurements*

Before doing the analysis all the wave forms were corrected for baseline EEG activity by subtracting the prestimulus electrical activity (for 50 ms before the presentation of stimulus). Table 3 shows, the latencies and amplitudes of peaks Pj, N_1, P_2 and N_2 for AN and normal hearing group. P_2/N_2 complex was present is all 14 individuals whereas P1/N1 complex was not present in 4 subjects. Whenever LLRs were present, latency and amplitudes were within normal range. Presence or absence of LLR peaks did not bear any relation to the speech identification scores. Pearson's product moment correlation failed to evidence any significant correlation between evoked potential parameters and other psychophysical test results. Table 4 shows latency, amplitude and area of MMN parameters. Area of the MMN was determined by calculating the area between wave and baseline and took into account both the duration and amplitude of MMN response. In 5 of 14 subjects, MMN could not be elicited. Pearson's product moment correlation was performed between MMN parameters and other psychophysical measures. Only peak latency of MMN evidenced a significant correlation with speech identification scores. As the number of subjects with MMN present was less, to interpret the results of correlation, a scatter plot was drawn between MMN peak latency and speech identification scores. As seen from the scatter plot (Figure 3), no trend could be observed between MMN peak latency and speech identification scores. Figure 4 shows the grand average of MMN waveform in AN subjects and normal hearing listeners. Whenever the MMN was present in individuals with AN, wave form was indistinguishable from normal listeners.

Table 3: Mean and SD (values in parenthesis) of amplitude and latencies of LLR components in both groups

		P.	N,	P,	$\overline{N,}$
AN subjects	Amplitude (in (J.V)	2.8 (0.9)	0.9 (0.8)	2.8 (2.08)	-I.I (2.3)
	Latency (in ms)	81 (16.2)	125.4(23.04)	154.1 (27.1)	205 (23)
Normal subjects	Amplitude (in u.V)	2.S (0.6)	-0.5 (0.5)	2.8(1.5)	-1.6(1.5)
	Latency (in ms)	69(15.2)	120.5(23.5)	145.3 (25.6)	200.2 (26.3)

Of the 9 subjects who had MMN, 5 of them could not behaviorally discriminate the two stimulus contrast (i.e. JND was more than 100 ms). Other 5 subjects who had no MMN also could not behaviorally discriminate the two-stimulus contrast. The MMN wave forms of those subjects who could behaviorally discriminate the stimulus contrast were virtually indistinguishable from those who did not behaviorally discriminate the contrast. This data indicate that presence of MMN does not necessarily indicate the presence of behavioral discrimination.

Discussion

The major findings of this research were:

i) Open set speech identification scores varied considerably in individuals with AN and speech identification scores had a significant correlation with the low frequency hearing sensitivity.

ii) All subjects with AN had severe temporal processing deficits as shown by JND measurements and TMTF.

iii) In majority of AN patients cortical evoked potentials could be recorded but none of the measured evoked potential parameters had any relation with psychophysical measurements.

Psychophysical measurements

Speech identification scores in AN individuals had good correlation with low frequency hearing sensitivity but not with the high frequency hearing sensitivity. This frequency specific correlation between hearing thresholds and speech identification scores, may be related to differential physiology between high frequency and low frequency coding. Low frequencies are usually coded by phase locked responses in type I auditory nerve fibers. Individuals with AN cannot use phase locking cues to the same extent as normal hearing listeners due to dyssynchronous firing of auditory nerve fibers. However, detection of the high frequency depends on place of excitation on basilar membrane and does not depend on the phase locking cues as much as low frequencies. We propose that, low frequency hearing sensitivity in these individuals may indicate the extent of temporal disruption in the auditory system. Its relation with speech identification scores is suggestive of importance of neural synchrony in understanding speech. This is also supported by other two observations:

i) A retrospective inspection of the data reveled, all 8 individuals who obtained speech identification scores more than 50% had their low frequency hearing sensitivity (average of 250 Hz, 500 Hz, and 1000 Hz) better than 25 dB HL and 6 individuals who had speech identification scores less than 20% had low frequency hearing sensitivity more than 40 dB HL.

ii) There was significant correlation between low frequency hearing sensitivity and peak modulation detection thresholds. Based on the above observations, we propose that low frequency hearing sensitivity in AN individuals may be the indicator of suprathreshold temporal processing deficits.

All AN individuals experienced severe difficulties in discriminating the speech stimuli that differed in time domain. As stimulus was presented at equal sensation levels to both the groups this resulted in difference in presentation levels (SPL) for each of the subjects. However, the difference in the JNDs for transition duration of syllable /da/ between two groups cannot be attributed to difference in presentation level (SPLs). It is shown that when the stimuli are sufficiently loud or at comfortable level auditory duration discrimination is independent of the intensity [14]. Individuals who had better discrimination abilities also possessed better open set speech identification scores. These findings stress the importance of perception of temporal variation in understanding speech information. Temporal processing deficits in individuals with AN are also demonstrated by poor performance on TMTF. Average peak sensitivity of individuals with AN was threefold more than the normals. Poor sensitivity to temporal modulations in these individuals is also reported by other investigators [6,8], A significant correlation was observed between modulation detection thresholds and speech identification scores (when data from two subjects with paradoxical results were removed). This finding agrees with the results obtained from Ranee et al [6], Zeng et al [7,8]

Table 4: Mean and SD (values in parenthesis) of amplitude and latencies of mismatch nagativity components in the auditory neuropathy group

	AN individuals			Normal subjects		
	MMN (On set)	MMN (Peak)	MMN (Off set)	MMN (On set)	MMN (Peak)	MMN (Off set)
Amplitude (in uV)	-0.068 (0.6)	-4.6(2.1)	1.9(3.2)	-0.071 (0.2)	-4.8(1.5)	1.5(2.5)
Latency (in ms)	117.3(23.6)	186.4(19.04)	209 (25.5)	120.5(20.5)	180.6 (20.8)	204 (25.5)

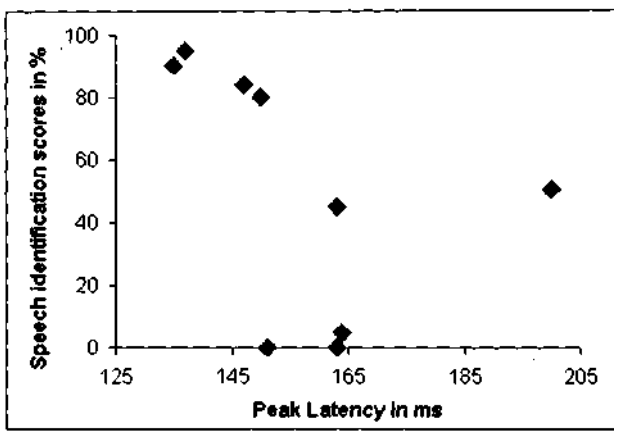


Figure 3
Scatter plot between speech identification scores and peak latency of MMN in auditory neuropathy subjects.

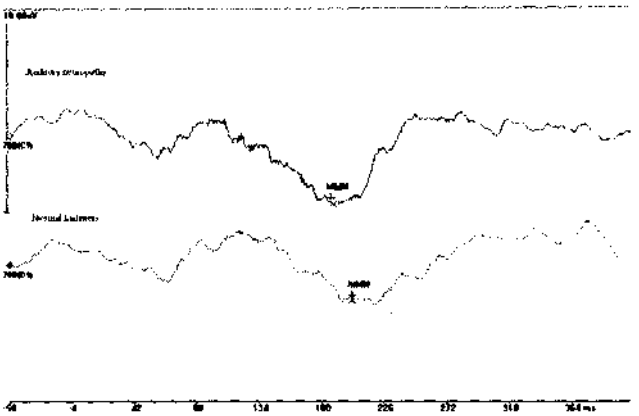


Figure 4
Grand averaged MMN wave form in the auditory neuropathy group and normally hearing subjects.

Difference between normal listeners and AN subjects in detection of modulation was more at higher modulation frequencies. The extent of temporal processing deficits were more than what is been reported for cochlear hearing loss of comparable degree [15]. This difference between two groups cannot be because of different presentation levels (SPA) used because modulation detection thresholds are reported to be stable over a wide range of intensities. In the auditory system, higher modulation frequencies are processed at auditory nerve and brainstem, whereas lower modulation frequencies are processed mainly in the thalamus and auditory cortex. As one ascends the auditory system, a neural encoding shift occurs. An emphasis on synchronous response for temporal coding exists at auditory nerve and brainstem (codes low frequencies) and less reliance on synchrony occurs as one move centrally (codes high frequencies) [16-18]. Hence, it can be expected that individuals with AN will have more problems in processing high rates of modulations which require synchronous firing of auditory nerve fibers. Inability of many the subjects to perceive amplitude modulation of 0 dB (100%) at higher modulation rates indicates the importance of temporal synchrony in auditory perception. Effects of reduced temporal fluctuations on speech perception in normal listeners have been reported previously [19]. Elevated modulation detection thresholds at slower modulation rates in combination with virtually no perception of modulations at high modulation rates are sufficient to disrupt the perception of amplitude envelop cues in normal speech. As this study measured only peak sensitivity, reduced peak sensitivity may also be due to reduced ability to perceive the amplitude changes in patients with auditory neuropathy.

Electrophysiological measures

P1/N1 and P_2/N_2 complex amplitude and latency did not appear to be related to degree of hearing loss or speech identification scores. This result is in contrast to Ranee et al [5] who evidenced a strong relation between presence of event related potential and speech perception scores. This difference in the results may be due to difference in subjects and the stimuli. Ranee et al [5] primarily studied children younger than 92 months and were fitted with the amplification devices before 28 months of age. This may have prevented the retrograde loss of speech perception abilities. In our subjects, average age at which amplification provided was 18 years. Many of the subjects were not identified in childhood as they had near normal hearing sensitivity and were grouped as slow learners in the class. This huge gap in the auditory experience between two groups might have adversely affected the speech perception abilities of the later. Presence of LLR components with normal latency and amplitude represent the stimulus registration in the primary auditory cortex, which do not involve complex decoding and representation of the signal as it is required for the speech perception.

Large numbers of studies in last decade have established MMN as an objective electrophysiological measure of auditory discrimination (e.g. [13]). Our results of MMN and behavioral discrimination are paradoxical. Significant MMN was seen in the majority of subjects with auditory neuropathy, even though stimulus contrast could not be behaviorally discriminated. Fried et al [20] have provided evidence for the existence of preconscious perception in the visual system. Preconscious perception describes the physiological or neurological process that occurs without behavioral or conscious perception. Some evidence of preconscious perception is also reported in auditory system using MMN. Allen et al. [16] reported the presence of

MMN in normal listeners for the stimulus contrast that they could not behaviorally discriminate. Presence of MMN in AN subjects who could not behaviorally discriminate the stimulus contrast supports the hypothesis that neural generators responsible for the MMN are not necessarily linked to conscious perception [21]. But all the individuals who had no MMN could not behaviorally discriminate the stimulus contrast. These two results in combination support the notion that MMN is necessary, but not a sufficient component for conscious perception of stimulus change.

Another possible explanation for the discrepancy between behavioral discrimination and MMN in some AN subjects may be related to perception of stimulus onset cues. We hypothesis that cues in the stimulus onset play a major role in the behavioral discrimination between the stimulus contrasts that differ in transition duration. Kraus et al [20] reported that perception of any change in the stimulus onset was extremely difficult in a subject with AN who had normal hearing. Hence the individuals with AN had larger JND's. We propose that MMN, which was present in some AN individuals, was elicited by the difference in the later part of the stimulus. However, it is unclear that why AN individuals could not discriminate the stimulus contrasts by using the information in the later part of the stimulus that elicited the MMN.

Conclusion

Findings of this study indicate that individuals with AN have severely affected temporal processing abilities. These temporal processing deficits correlate significantly with the speech identification scores and hearing sensitivity in the low frequency region. Psychophysical measures including speech perception did not correlate with the electrophysiological measurements used at least in adults with AN.

Authors' contributions

AKU was involved in designing the study, data collection, analysis, interpretation and preparing the manuscript. JM was involved in designing die study, interpretation and preparing the manuscript.

References

1. Starr A, Picton TW, Sininger Y, Hood **L,** Berlin C: **Auditory neuropathy.** *Brain* 1996, **119:741-753.**
2. Salvi RJ, Wang **J,** Ding D, Stecker N, Arnold S: **Auditory deprivation in the central auditory system resulting from selective inner hair cell loss: Animal model of auditory neuropathy.** *Scand Audiol* **1**999, **51 (Suppl l):** I-12.
3. Starr A, Michalewski HJ, Zeng FG, Brooks SF, Linthicum F, Kim CS, Winnier D, Keats B: **Pathology and physiology of auditory neuropathy with a novel mutation in the MPZ gene,** *brain* 2003, **126:1604-1619.**
4. Starr A, Sininger YS, Praat **H: The varieties of auditory neuropathy.** *J Basic Clin Physiol Pharmacol* 2000, **I** 1:215-229.
5. Ranee G, Cone-Wesson B, Wunderlich J, Dowell R: **Speech perception and cortical event related potentials in children with auditory neuropathy.** *Ear Hear* **2002, 23:239-253.**
6. Ranee G, McKay C, Grayden D: **Perceptual characterization of children with auditory neuropathy.** *Ear Hear* 2004, **25:34-46.**
7. Zeng FG, Kong YY, Michalewsk HJ, Starr A: **Perceptual consequences of disrupted auditory nerve activity. J** *Neurophysiol* 2005, 93:3050-3063.
8. Zeng FG, Oba S, Garde S, Sininger Y, Starr A: **Temporal and speech processing deficits in auditory neuropathy.** *Neuroreport* 1999, **10:3429-3435.**
9. Vandana : **Speech identification test in Kannada.** *Independent project submitted to University of Mysore, Mysore* 1998.
10. Babu RM, Ratna N, Bettagiri R: **Test of articulation in Kannada.** *ThejAIISH* 1972,3:7-19.
11. Moulines E, Laroche J: **Non parametric techniques for pitch scale and time scale modification of speech.** *Speech Commun* **1995, 16:175-205.**
12. Taylor MM, Creelman CD: **PEST: Efficient estimates on probability functions.** *J Acoust Soc Am* 1967, 41:782-787.
13. Naatanen R: **The mismatch negativity: A power full tool for cognitive neuroscience.** *Era Hear* **1995, 16:6-18.**
14. Moor BCJ, Shailer MJ, Schooneveldt GP: **Temporal modulation transfer function for band noise in subjects with cochlear hearing loss.** *Br]* Audio/ 1992,26:229-237.
15. Creelman DC: **Human discrimination of auditory duration. J** *Acoust Soc Am* **1962, 34:582-593.**
16. Wang X, Sachs MB: **Neural encoding of the single formant stimuli in cat. I Responses of anteroventral cochlear neurons.** *J Neurophysiol* **1993, 71:59-78.**
17. Allen J, Kruas, Bradlow Wang X, Sachs MB: **Neural representation of consciously imperceptible speech sound differences.** *Percept Psychophys* 2000, **62:**1383-1393.
18. Frisina RD: **Subcortical neural coding mechanisms for auditory temporal processing.** *Hear Res* **200**1, 1**58:**1-27.
19. Drullman R, FestenJM, Plomp R: **Effects of temporal smearing on speech perception.** *J Acoust Soc Am* **1994, 95:1053-1064.**
20. Fried **I,** MacDonald KA, Wilson CL: **Single neuron activity in human hippocampus and amygdala during the recognition of faces and objects.** Neuron **1997,** 18:753-765.
21. Kraus N, Bradlow AR, Cheatham MA, Cunningham J, King CD, Koch DB, Nicol TG, McGee TJ, Stein LK, Wright **BA: Consequences of neural asynchrony: A case of auditory neuropathy.** *JAssoc* Res *Otolaryngol* 2000, 1:33-45.
22. Naatanen R: **Phoneme representations of the human brain as reflected by event-related potentials.** *Electroencephalogr Clin Neurophysiol Suppl* 1999, **49:170-173.**
23. Vandana : **Speech identification test in Kannada.** *Independent project submitted to University of Mysore, Mysore* **1998.**

anuscript (include abstract, figure captions, references)

Processing of Temporal Modulations in a Digital Hearing Aid

ABSTRACT

The purpose of the present study was to investigate the processing of temporal envelope in a digital hearing aid. Perception of temporal modulations through a digital hearing aid was also investigated under linear and compression processing strategies in a hearing aid. Modulation detection thresholds were measured for sinusoidally amplitude modulated white noise at 4 Hz, 16 Hz, 50 Hz, and 100 Hz in 10 normal hearing adults for three conditions of (a) audiometer-loudspeaker output, (b) audiometer-hearing aid-loudspeaker output under compression mode, and (c) audiometer-hearing aid-loudspeaker output under linear mode. In an another experiment, the input and output functions of the test hearing aid was compared for sinusoidal amplitude modulated white noise at these modulation frequencies with 0 dB (100%) amplitude modulation. Results showed that the hearing aid employed in this study did not influence the modulation detection thresholds for low modulation frequencies, but it increased detection thresholds for 100 Hz. The results also showed that the hearing aid did not replicate the temporal modulations of the original signal at any of the modulation frequencies in either compression or linear mode. These results suggest that it is not the processing strategy (compression or linear), per se, that influences perception of temporal information. Perhaps, it is the distortion induced by hearing aid while processing rapid temporal changes which interferes with speech perception. The results of the present study

suggest a need to revisit signal-processing strategies, particularly as related to temporal information of speech, in digital hearing aids.

Processing of Temporal Modulations in a Digital Hearing Aid

The temporal envelope of a signal can be thought of as the line connecting successive peaks in a stimulus waveform (Scharf and Buss, 1986). In complex signals, such as speech and music, spectral composition changes as a function of time. In other words, speech signal may be described as a summation of amplitude modulated narrow frequency bands. From this perspective, it can be said that every frequency band of a speech signal has a carrier (fine structure) as well as a modulated component. These modulations are significant for the perception of speech sounds (Drullman, Festen and Plomp, 1994). Drullman, Festen and Plomp (1994) studied the effect of smearing of temporal envelope on sentence intelligibility and phoneme recognition. They divided the wide band speech signal into a series of frequency bands. Amplitude envelope of each of these frequency bands was high pass filtered. They presented speech, modulated with this high pass filtered envelope, to 36 normal hearing individuals. The results indicated that speech recognition significantly reduced when temporal envelope was high pass filtered above 64 Hz. They also reported that temporal smearing of signals affected perception of consonants (especially the stops) more than that of vowels. It is thus clear that hearing aids should preserve low frequency modulation below 64 Hz while processing speech. Failure to do so will result in temporal smearing of the signal affecting speech intelligibility. In fact, near perfect speech recognition can be obtained even if spectral information is greatly reduced, but if temporal envelope cues are preserved in at least a few frequency bands (Shannon, Zeng, Kamath, Wygnoski and Ekelid, 1995). Therefore, speech processing and

transmitting strategies in hearing aids should aim at faithful preservation of low frequency temporal modulations.

Modification of temporal information during amplification and its effect on the intelligibility of speech has been a topic of discussion for many years. Some researchers have maintained that multi-band amplitude compression circuits, particularly those with short attack and release times, adversely affect speech recognition by reducing the natural temporal fluctuations of speech (Plomp, 1988; Moore, Johnson, Clark and Pluvinage, 1992). Boothroyd and Nittrouer (1988) reported that performance of their subjects on a phoneme discrimination task reduced when compression was introduced. Perhaps, amplitude compression reduces both temporal and spectral contrast in the speech signal (Plomp, 1988).

One way of describing temporal resolution of a given system is to measure temporal modulation transfer functions (TMTFs). TMTF represents detection thresholds for amplitude modulated signals as a function of modulation rate (/m). This approach has extensively been employed in psychophysical studies of auditory temporal resolution in individuals with normal hearing (eg: Viemeister, 1979), with sensori-neural hearing loss (Bacon and Viemeister, 1985; Moore and Glasberg, 2001; Moore, Shailer, Schooneveldt, 1992), with cochlear and brainstem implants (Shannon, 1992; Busby, Tong and Clark, 1993), with cortical damage (Lorenzi, Wable, Moroni, Derobet and Frachet, 2000), and in individuals with developmental dyslexia (Lorenzi, Dumont and Fullgrabe, 2000). A strong relationship has been reported between modulation detection thresholds and speech recognition/phoneme identification scores in individuals with cochlear implants

(Cazals, Pelizzone, Saudan and Boex, 1994), and in patients with auditory neuropathy (Kumar and Jayaram, 2005; Ranee, McKay and Grayden, 2004; Zeng, Kong, Mechalewski and Starr, 2005).

As temporal envelope of the signal is very important for the perception of signals, the present study investigated if temporal modulations of the input signal are preserved following processing of these signals in a digital hearing aid. Specifically, processing of temporal modulations through linear and compression algorithms of a digital hearing aid was compared. In other words, the objective of the present study was to see if a digital hearing aid induces distortions while processing temporal modulations at different modulation frequencies. This was done by measuring TMTFs, in normal listeners, in three conditions: without hearing aid, through hearing aid programmed for linear and compression strategies. Processing of temporal modulation in a digital hearing aid was objectively assessed by comparing the input and output of the hearing aid under linear and compression conditions.

METHOD

Subjects

Ten normal hearing listeners, all males, in the age range of 18 -22 years (mean age 20 years) participated in the study. All subjects had hearing thresholds of < 15 dB HL in octave frequencies between 250 to 8000 Hz and had normal results on immittance evaluation. None of the subjects reported any history of ear disease, or exposure to loud noise, or usage of ototoxic drug.

Hearing aid details

Pbonak Supero, a commercially available 5-channel digital hearing aid, was used in the study. In the compression mode, the hearing aid employed a compression ratio of 1: 2.3 across the frequency range with an attack time of 5ms and a release time of 8ms. Threshold knee point was kept at 40 dB to ensure that the test signals were all above the knee point. The hearing aid amplified the signal with a compression ratio of 1:1 over a range of 80 dB SPL in the linear mode

Stimuli

The stimuli consisted of unmodulated and sinusoidally amplitude modulated broadband noise of 500ms with a ramp of 2.5ms. The modulated signal was derived by multiplying the broadband noise by a dc-shifted sine wave. The depth of modulation was controlled by varying the amplitude of modulating sine wave. Modulation depth for different stimuli varied between 0 to -30 dB (where 0 dB is equal to 100% modulation depth) with a step size of 3 dB. Detection thresholds were measured for 4 modulated frequencies: 4 Hz, 16 Hz, 50 Hz and 100 Hz. All the stimuli were generated using a 32 bit D/A converter at a sampling frequency of 44.1 kHz. Figure 1 shows one of the stimuli used in the study (4 Hz modulation at 100% and 50% modulation depth).

Insert Figure 1 about here

Procedure

Experiment 1

This experiment was designed to measure modulation detection thresholds for different modulation frequencies. Sinusoidally amplitude modulated white noise at different modulation frequencies was played through a personal computer at a sampling frequency of 44.1 kHz which was later inputted into an audiometer (Maico MA-53). Subjects received the stimuli through a loudspeaker connected to the audiometer. Loudspeaker was positioned at a distance of 1 meter and at 90° azimuth (Figure 2a).

The worst threshold that can be obtained corresponds to a modulation depth of 0 dB (100% modulated noise). Intensity of the stimulus was kept at 40 dB SL (this value was always below 60 dB SPL) in all the conditions. The output of the loudspeaker was calibrated using Quest 1800 sound level meter and 4180 free field microphone in the beginning of the experiment. Selection of the ear was done randomly. Non-test ear was occluded with E A R (Etymotic Research, U S A) earplugs.

Insert Figure 2 about here

Modulation detection thresholds were determined using AX 'same-difference' discrimination paradigm. In this paradigm, A is the anchor stimulus while X is the variable stimulus. The subjects' task was to indicate whether A is the same as X or not. Anchor stimulus was always unmodulated white noise while variable stimulus was sinusoidally amplitude modulated noise. The depth of amplitude modulation was varied according to the following rule: reduce the depth by 3 dB following two correct responses, and increase it by 3 dB following an incorrect response. This procedure provides an estimate of the value necessary for 70.7% correct response (Levite, 1979). Inter- stimulus

interval between anchor and variable stimulus was 500ms. Modulation detection thresholds were measured, with the procedure described here, for four modulation frequencies: 4 Hz, 16 Hz, 50 Hz and 100 Hz. Test trials also included catch trials. A catch trial consisted of either two identical anchor or two identical non-anchor stimuli. Subjects were tested individually in a sound treated room.

Experiment 2

The first experiment was repeated, but this time, the output from the audiometer was fed to the loudspeaker through the test hearing aid (Phonak supero). Furthermore, the hearing aid was programmed for either compression or linear algorithm. The output from the audiometer was routed through the hearing aid to the loudspeaker (see Figure 2b). The experimental procedure was otherwise the same as in experiment 1. The order of presentation of modulation frequencies and conditions was randomised to nullify the order and learning effect.

Experiment 3

The temporal envelope of the input and output of the test hearing aid was analysed. A block diagram of the experimental setting is shown in Figure 3. The test hearing aid was kept in an anechoic box and the input stimulus was sinusoidally amplitude modulated white noise (4 Hz, 16 Hz, 50 Hz and 100 Hz). A microphone (mid), kept near the microphone port of the hearing aid, picked up the input to the hearing aid and stored it in a computer at a sampling frequency of 44.1 kHz. Output of the hearing aid was routed, through a 2cc coupler and a second microphone (mic2), to a computer which stored this signal at a sampling rate of 44.1 kHz. Both the microphones (Maico 188) had similar

response characteristics. A correlation analysis was performed between the temporal envelope of the input and output of the hearing aid using MATLAB software. This procedure was repeated under two conditions: with and without the compression circuit on.

Insert Figure 3 about here

RESULTS

Modulation Detection Thresholds

Table 1 shows the mean and standard deviation of modulation detection thresholds across different frequencies for the three conditions tested (results of experiments 1 and 2). Multivariate analysis of variance was used to test the significance of difference between the modulation detection thresholds obtained with and without the hearing aid. MANOVA revealed a significant main effect of hearing aid on modulation detection thresholds [$F(4,16) = 5.41$, $p < 0.05$].

Insert Table 1 about here

Bonferroni's post-hoc comparison revealed that modulation detection thresholds obtained without the hearing aid differed significantly from those obtained under both linear and compression conditions at 100 Hz. However, the modulation detection thresholds did not differ significantly between linear and compression conditions themselves. Figure 4 shows error bars for modulation detection thresholds for the three conditions: without hearing aid, linear and compression condition of hearing aid.

Insert Figure 4 about here

Input-Output of Hearing Aid

Table 2 shows correlation coefficients between the temporal envelope of input waveform and output of the hearing aid for linear and compression conditions of the hearing aid. If the hearing aid replicated the temporal

Insert Table 2 about here

envelope of the signal without any distortion, then the correlation coefficient between input and output would be 1. Thus, any value closer to 1 indicates good replication of the input by the hearing aid while values away from 1 reflect temporal distortion induced by the hearing aid. In other words, the lower the correlation coefficient (nearer to 0), the greater is the distortion introduced by the hearing aid. As can be seen from Table 2, there was a significant correlation between the input and output of the hearing aid only in 'linear' condition and at the modulation frequency of 4 Hz. This shows that the hearing aid did not reflect the temporal changes of the input waveform at any of the other modulation frequencies. Figure 5 shows a representative time domain wave form of input to, and output from, the hearing aid. Temporal distortion induced by the hearing aid was more in 'compression' condition than in 'linear' condition, and it increased with modulation frequency (interested readers can get raw data from the authors).

Insert Figure 5 about here

Discussion

Modulation detection thresholds did not differ significantly between unaided and aided conditions, except for 100 Hz modulation frequency. Modulation detection threshold was significantly higher with the hearing aid than without the hearing aid at this modulation frequency. However, the signal processing strategy (linear vs. compression) did not influence modulation detection thresholds. These results, in combination, suggest that it is not compression which is influencing the detection thresholds of modulations, but, on the other hand, it is the failure of the hearing aid to reflect the rapid temporal fluctuations that is responsible/or f&e increase m mocWafz'oM ckfecfzoM fAr&sWd. Previous research (Plomp, 1988) has shown that multi-channel amplitude compression employed in digital hearing aids reduces the 'modulation-transfer function', thereby reducing the intelligibility of speech. The 'modulation transfer function' describes the extent to which speech intensity envelope is transferred by the device. Modulation transfer function is equal to the modulation index as a function of frequency at the output of the transmission channel for 0 dB (100%) amplitude modulated input signal. Though multi-channel compression has been shown to affect modulation transfer function (Plomp, 1988), it does not seem to have affected modulation detection thresholds in the subjects of the present study. All the subjects in the present study had modulation detection thresholds of less than -11.5 dB. Perhaps, such a small change in intensity over time caused the hearing aid to operate in either compression or linear mode, as the case may be, throughout the modulation cycle and did not allow the hearing aid to toggle between compression and linear modes, thereby reducing distortion.

Compression may have a detrimental effect on signal processing and perception of high modulation depths. This was evident from the results of experiment 3. When the input to the hearing aid was 0 dB (100%) amplitude modulated noise, hearing aid did not replicate the temporal envelope of the signal at any of the modulation frequencies. It is clear from the correlation coefficients that distortion of temporal envelope increased with modulation frequency. The results of experiment 3 on simultaneous measurement of the input to, and the output of, hearing aid using two identical microphones rule out the possibility of the signal getting distorted at the signal acquisition stage. The distortion may be due to some amount of ringing of the diaphragm of the hearing aid microphone which might have interfered with the processing of modulations. Another factor that might have contributed to distortion of temporal envelope was the response of hearing aid microphones to starting transients. This is the ability of diaphragm of the microphone to respond rapidly to the changes that occur at the start of the sound (Talbot-Smith, 2001). In the 0 dB (100%) amplitude modulated noise, response of the microphone to the start of every modulation cycle can be considered as starting response and the inability of the microphone diaphragm to respond to this rapid temporal change may have induced distortion in the processing of the signal.

Speech contains both spectral (formant frequencies and transitions) and temporal cues (fine structure and modulating envelope). Many studies have shown that temporal cues may, by themselves, lead to good recognition of consonants, vowels and sentences (Souza & Kitch, 2001; Shannon et al., 1995, among others). Temporal cues must be preserved for faithful transfer of speech signal. It can be hypothesised that a hearing aid which attenuates high, rather

than low, frequency components of the envelope to a greater extent may reduce the contrast between consonants in the process. This may, in turn, reduce the overall speech intelligibility. Digital technology has enhanced the use of speech processing strategies in hearing aids. This has led to development of newer strategies for noise reduction, spectral enhancement, and an increase in the number of channels in hearing aids. Even with all these advancements, it has been shown (Weigand, Bodkin, Madell, Rosenfeld, & Press, 2002) that around 9% of the hearing aid users return their hearing aids and that a majority of them do so because of what they consider as 'insufficient benefit' from the hearing aids. The perception of 'insufficient benefit' may have something to do with the 'inadequate' processing of rapid temporal changes of the stimulus in the hearing aids resulting in less intelligible and unnatural speech. The results of the present study suggest a need to revisit signal processing strategies, particularly as related to temporal information of speech, in digital hearing aids.

In general, the data of the present study showed that the hearing aid studied did not alter modulation detection thresholds at low modulation frequencies, but it significantly increased modulation detection thresholds at high modulation frequencies. Besides, the test hearing aid induced significant level of temporal distortion at all modulation frequencies. Temporal distortion indicates poor processing of rapid changes in amplitude by the hearing aid. Effects of reduced perception of modulations are well known. Smearing of the temporal envelope significantly reduces sentence intelligibility and phoneme recognition (Drullman, Festen and Plomp, 1994). Poor speech perception in individuals with auditory neuropathy has been related to poor modulation detection thresholds (Kumar and Jayaram, 2005). These patients generally do

not benefit from hearing aids probably because hearing aids are not sufficiently responsive to temporal modulations in the input signal, and thus, may induce significant temporal distortion. Therefore, technology should focus on reducing temporal distortion in the hearing aid to facilitate better speech perception for individuals with hearing impairment.

REFERENCES

Bacon, S.P., & Veimeister, N.F. (1985). Temporal modulation transfer function in normal hearing and hearing impaired subjects. Audiology, 24,117-134.

Boothroyd, A., & Nittrouer, S. (1988). Mathematical treatment of context effects in phoneme and word recognition. Journal of the Acoustical Society of America, 84,101-114.

Busby, P.A., Tong, Y.C., & Clark, G.M. (1993). The perception of temporal modulations by cochlear implant patients. Journal of the Acoustical Society of America, 94,124-131.

Cazals, Y., Pelizzone, M., Saudan, O., & Boex, C. (1994). Low pass filtering in amplitude modulation detection associated with vowel and consonant identification in subjects with cochlear implants. Journal of the Acoustical Society of America, 96,2048-2054.

Drullman, R., Festen, J.M., & Plomp, R. (1994). Effect of temporal envelope smearing on speech perception. Journal of the Acoustical Society of America, 95,1053-1064.

Kumar U.A., & Jayaram, M. (2005). Auditory processing in individuals with auditory neuropathy. Manuscript accepted in Behavioural Brain Functions.

Levett, H. (1971). Transformed up-down methods in psychoacoustics. Journal of the acoustical society of America 49, 467-477.

Lorenzi, C, Dumont, A., Fullgrabe, C. (2000). Use of temporal envelope cues by developmental dyslexics. Journal of Speech, Language and Hearing Research, 43, 1367-1379.

Lorenzi, C, Wable, 1, Moroni. C, Derobet, C, Frachet, B. (2000). Auditory temporal envelope processing in a patient with left hemisphere damage. Neurocase 6: 231-244.

Moore, B.C.J., & Glasberg, B.R. (2001). Temporal modulation transfer functions obtained using sinusoidal carriers with normal hearing and hearing impaired listeners. Journal of the Acoustical Society of America, 110, 1067-1073.

Moore, B.C.J., Johnson, J.S., Clark, T.M., & Pluvinage, V. (1992). Evaluation of a dual channel full dynamic range compression system for people with sensory neural hearing loss. Ear and Hearing, 13, 349-370.

Moore, B.C.J., Shailer, M.J., & Schooneveldt, G.P. (1992). Temporal modulation transfer function for band noise in subjects with cochlear hearing loss. British Journal of Audiology, 26, 229-237.

Plomp, R. (1988). The negative effect of amplitude compression in multichannel hearing aids in the light if modulation transfer function. Journal of the Acoustical Society of America, 83, 2322-2327.

Ranee, G., McKay, C, & Grayden, D. (2004). Perceptual characterization of children with auditory neuropathy. Ear and Hearing, 25, 34-46.

Scharf, B., & Buss S. (2000). Audition. 1. Stimulus, physiology, thresholds. In K. R. Boff, L. Kaufman, & J.P. Thomas (Eds). Handbook of perception

and human performance: Sensory process and perception (pp, 141-147). New York: Wiley.

Shannon, R.V. (1992). Temporal modulation transfer functions in patients with cochlear implants. Journal of the Acoustical Society of America, 92, 2156-2164.

Shannon, R.V., Zeng, F.G., Kamath, V., Wygnoski, J., & Ekelid, M. Speech perception with primarily temporal cues. Science 1995; 270; 303-304.

Souza, P.E., & Kitch, V. (2001). The contribution of amplitude envelope cues to sentence identification in young and aged listeners. Ear and Hearing, 22, 112-119.

Talbot-Smith, M. (2001). Audio engineers reference book. New York: Focal Press.

Viemeister, N.F. (1979). Temporal modulation transfer functions based on modulation detection thresholds. Journal of the Acoustical Society of America, 66, 1364-1380.

Reference Note

Weigand, J., Bodkin, K., Madell, J.R., Rosenfeld, R.M., & Press, I.L. (2002). By the numbers: Stastical analysis of hearing instrument returns. Hearing Review, 9 (1).

Zeng, F.G., Kong, Y.Y., Mechalewski, H.J., & Starr, A. (2005). Perceptual consequences of disrupted auditory nerve activity. Neurophysiology, 93, 3050-3063.

Table 1. Mean modulation detection thresholds, and standard deviations (SD), under three conditions

	Unaided		Linear		Compression	
	Mean (dB)	SD	Mean (dB)	SD	Mean (dB)	SD
4 Hz	-18.01	2.7	-17.3	4.2	-14.6	2.4
16 Hz	-15.1	4.3	-13.3	6.00	-11.9	4.5
50 Hz	-13.0	2.8	-12.0	3.4	-11.6	5.8
100 Hz	-10.4	1.4	-8.5	1.2	-8.0	.7

Table 2 Correlation coefficients between input and output of the hearing aid for different modulation frequencies

	4 Hz	16 Hz	50 Hz	100 Hz
Linear	0.6*	0.025	0.053	-0.132
compression	0.016	0.007	0.001	0.004

* $p < 0.05$

Legends for figures

1. Figure 1. (a) Unmodulated, (b) 100% amplitude modulated noise at 16 Hz and (c) 50% amplitude modulated noise at 16 Hz.
2. Figure 2. (a) Block diagram of the settings for experiment 1 and (b) experiment 2
3. Figure 3. Block diagram of the settings for Experiment 3.
4. Figure 4. Modulation detection thresholds for different conditions (error bars show mean + 1 SD).
5. Figure 5. (a) Input, and output of the hearing aid in (b) linear and (c) compression modes for 100% amplitude modulated white noise at 16 Hz.

Figure
Click here to download high resolution image

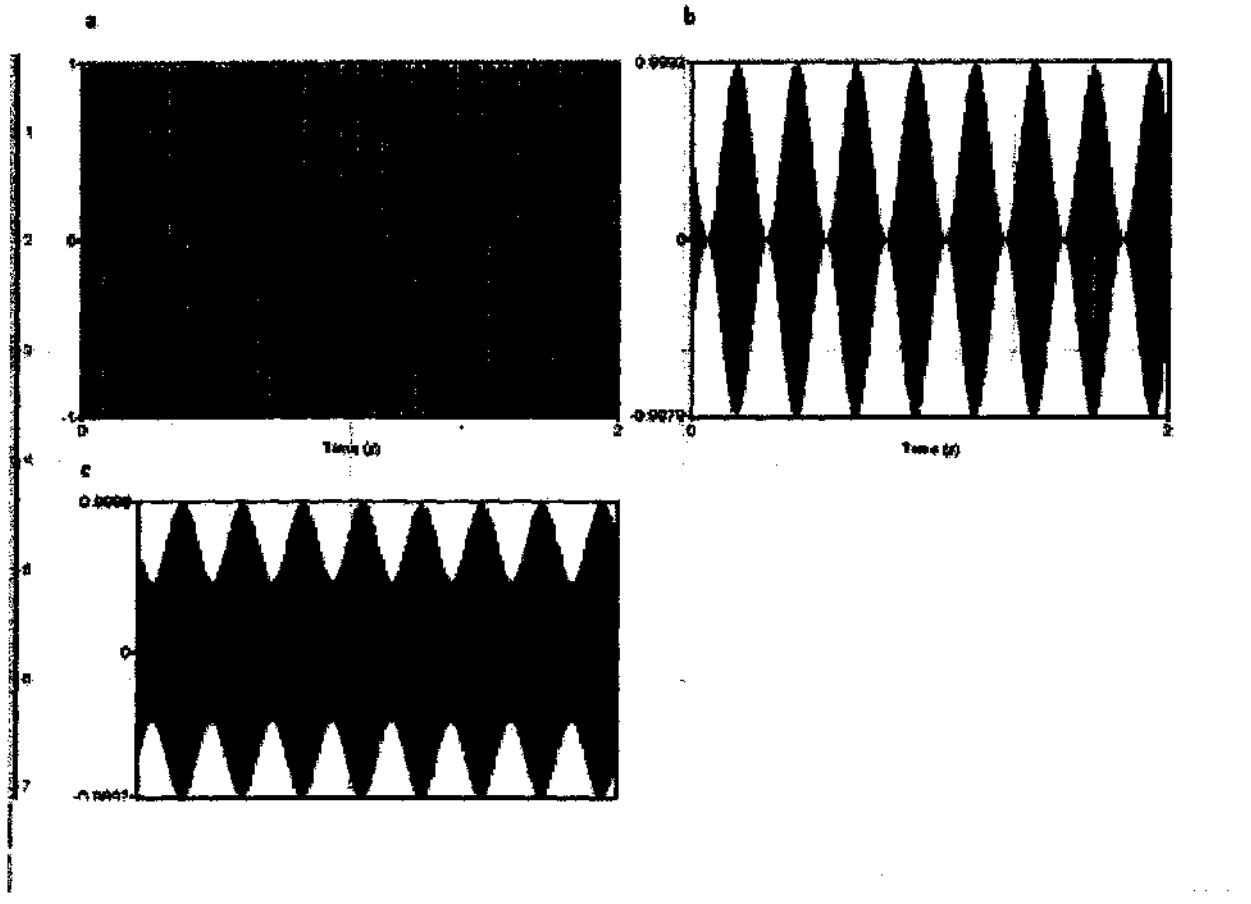


Figure
Click here to download high resolution image

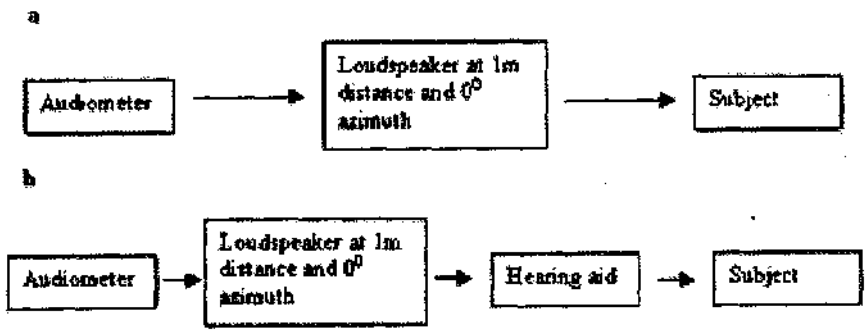


Figure
Click here to download high resolution image

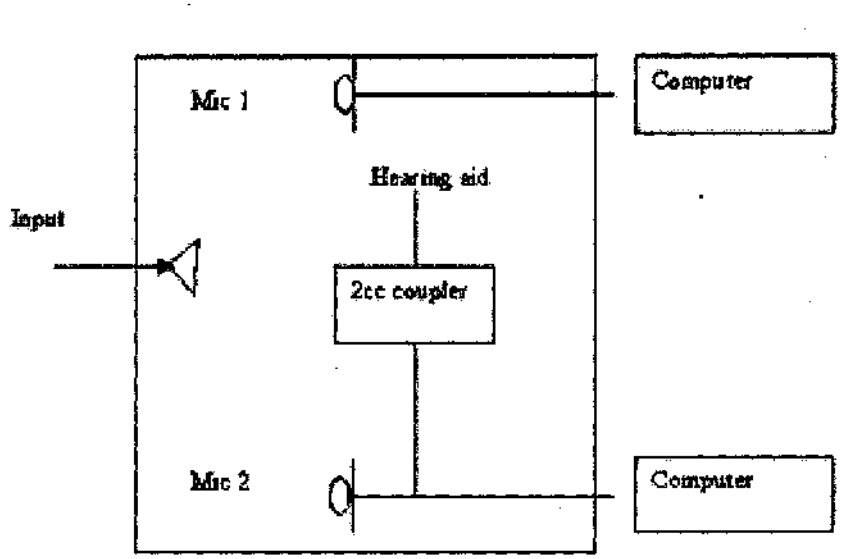


Figure
Click here to download high resolution image

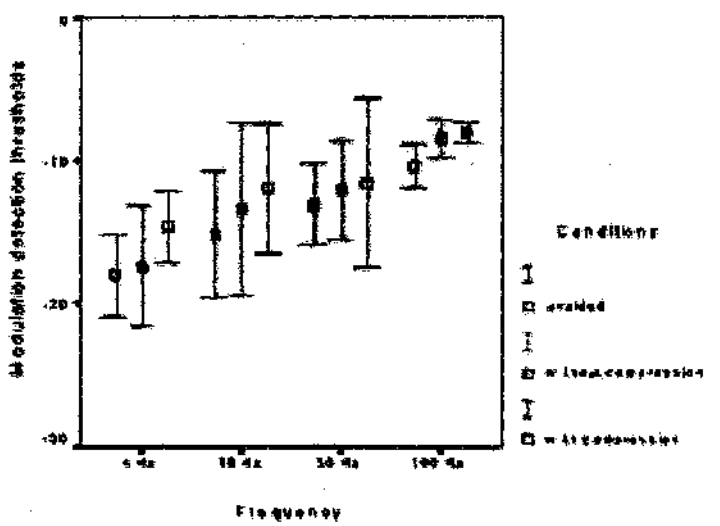
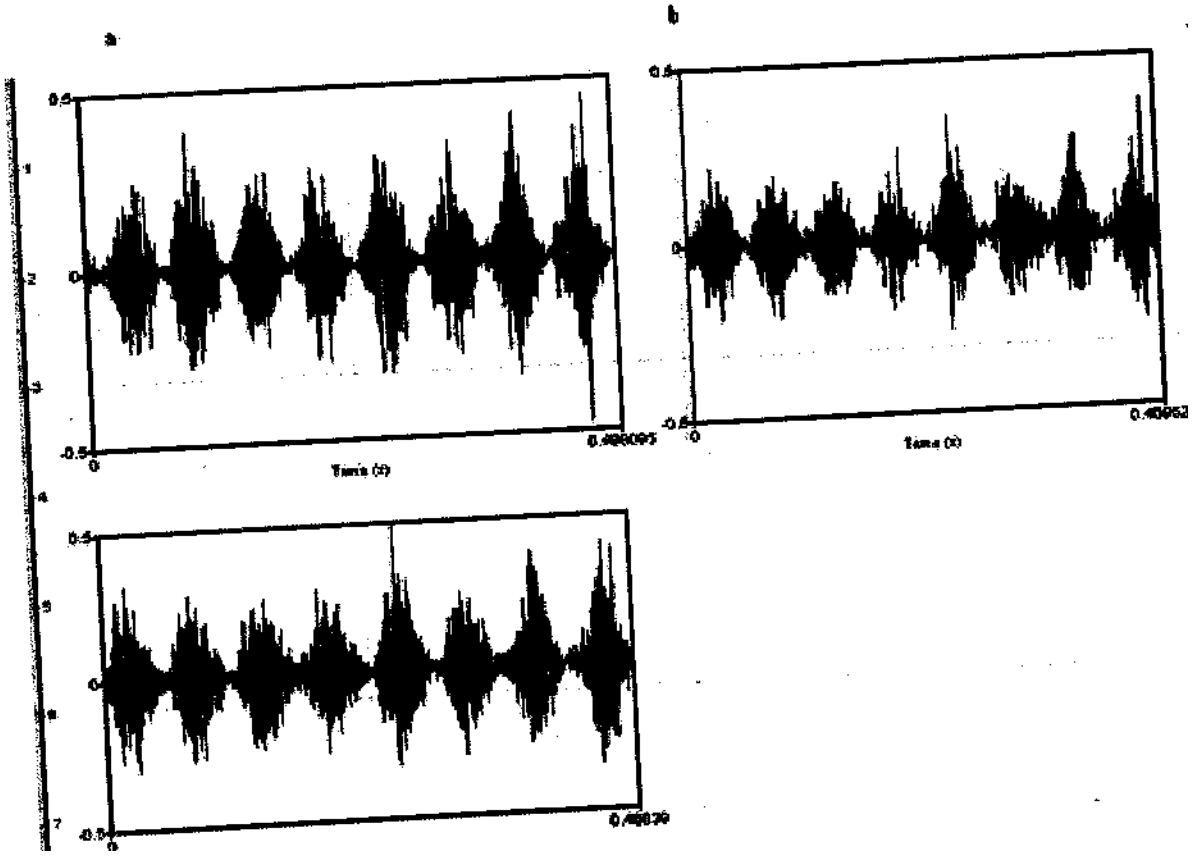


Figure
Click here to download high resolution Image



PRESENTATION /INVITED TALKS

Kumar, U. A. (2005). 'Speech Perception - Overview' in the National Seminar on Phonetic Perception held at AIISH, Mysore, on 25th to 26th August, 2005.

Kumar, U.A. (2004). 'Speech processing through digital hearing aids-requirements' in the National Seminar on Speech processing held at AIISH, Mysore, on 4th to 5th November, 2004.

Kumar, U.A. (2004). Tsychophysics and speech perception abilities in individuals with auditory dys-synchrony' in the Seminar on auditory dys-synchrony held at AIISH, Mysore, on 22nd September, 2004.

Kumar, U.A., and Jayaram, M. (2005). Incidence prevalence and Audiological characteristics of auditory neuropathy. Paper presented at 37th national conference of India Speech and Hearing Association January, 2005, Indoor, India. Awarded best paper in Audiology.

Kumar, U.A., & Jayaram, M. (2005). Processing and perception of temporal modulations through digitl hearing aid. Paper presented at meeting of Acoustical Society of India, November, Mysore, India. Awarded best paper in Speech Communication