AUDIOVISUAL PERCEPTION AND PROCESSING IN INDIVIDUALS WITH AUDITORY DYSSYNCHRONY

AIISH Research Fund Project No. SH/CDN/ARF/3.87/2010-11

August 7, 2010 to September 7, 2011

Principal Investigator

Sandeep M., Ph.D. (Speech and Hearing)

Co-Investigator

Geetha C., M.Sc. (Audiology)

AIISH Research Fund Project

Sanction No. SH/CDN/ARF/3.87/2010-11

Personnel	Dhanya M Research Officer
Total Fund	Rs. 3,11,000/-
Project duration	13 months
Principal Investigator	Dr. Sandeep M Lecturer, Dept. of Audiology All India Institute of Speech and Hearing Manasagangothri, Mysore – 570 006, Karnataka, India
Co-Investigator:	Ms. Geetha C Lecturer in Audiology, Department of Audiology All India Institute of Speech and Hearing Mysore– 570 006, Karnataka, India

Acknowledgments

The investigators would like to acknowledge the Director of the All India Institute of Speech and Hearing for being pivotal in granting the funds for the study. We would like to acknowledge the HOD, Department of Audiology for allowing us to use the resources from the department for testing. We also would like to extend our appreciation to the accounts section for providing support in maintaining the accounts. Our heartfelt gratitude is also extended to all the participants for their kind cooperation. Lastly we would like to express our thanks to Nike Gnanateja for his timely help.

Contents

List of Tables

Abstract

There has been considerable increase in the prevalence of Auditory neuropathy spectrum disorders (ANSD). It is a known fact theat the benefit of listening devices in Individuals with ANSD is very limited and these individuals are often advised to practice speech reading along with the use of listening devices. However, little is known about whether they combine or integrate auditory and visual information in the normal way or not. Hence, the primary objective of the present study is to document audiovisual perception in individuals with ANSD. The study included 82 adults (48 males & 32 females) in the age range of 12 to 50 years. Of the 82 participants 42 (clinical group) were individuals with ANSD and 40 (control group) were normal hearing individuals. Audio and video recording of CV syllables /ba/, /da/ and /ga/ were done to create auditory (AO), visual (VO), audiovisual congruent (AVc), and audiovisual incongruent (AVi) conditions. Behavioral speech identification was done and auditory late latency responses (ALLR) were recorded using these stimuli. The results of behavioral speech identification revealed that ANSD group performed poorer than the control group in all the stimulus conditions for all the syllables. However, the performance of both the groups varied depending on the place of articulation and stimulus condition. Sequential information transfer analysis (SINFA) was performed. Results of this revealed that AO and AVi conditions yielded poorer performance with least performance in AVi condition. The maximum difference between the two groups was evident in AO condition. The ALLR waves were analyzed for the peak latency, amplitude and scalp topography. The mean latencies were prolonged in ANSD group compared to control group. When elicited by /da/ and /ga/, the latencies were earlier in the AVi condition than AO and AVc conditions. Whereas no such trend was seen with /ba/ as the stimulus. Topographic maps showed differences between the two groups and also across stimulus conditions.

Chapter 1

INTRODUCTION

Speech is the most common mode of communication among human beings and auditory system plays an important role in the speech perception. Through the auditory system, one is able to get not just the necessary cues but also the redundant cues of speech. Although normal hearing individuals in ideal listening conditions are able to get essential as well as redundant cues, individuals with hearing impairment are likely to miss out on even the essential cues. Therefore, the enhancement of speech perception is one of the key factors in the successful management of individuals with hearing impairment. The classical approach in this direction has been utilizing and integrating the cues from auditory and visual domains.

Face-to-face communication is more effective than situations involving just the voice, especially in a noisy environment (Sumby & Pollack, 1954; Benoit, Mohamadi & Kandel, 1994). As stated by Rosenblum (2005), "human speech is a multimodal function, usually apprehended by visual (speech reading) as well as auditory (hearing) means". The term ‘‘auditory-visual integration’’ is used to denote the processes employed by individual receivers to combine the information extracted from Auditory and Visual sources (Massaro, 1998).

The auditory signal provides three main cues for identifying consonants. These include place of articulation, manner of articulation and voicing. Place of articulation refers to where in the mouth the sound is produced, or the location of constriction in the oral cavity. These locations consist of bilabials, labiodentals, interdentals, alveolars, palatal-alveolars, palatals and velars. The manner of articulation describes how the articulators come in contact with one another to form a sound. Manners include stops, fricatives, affricates, liquids and glides. The third cue is voicing, which refers to the presence or absence of vocal fold vibrations. All of this information is incorporated in the spectral and temporal envelopes of the speech waveform (Ladefoged, 2006).

As shown by McGurk and McDonald, visual inputs also serve as important cues for speech perception. However, less information can be obtained from visual cues. Place of articulation is essentially the only observable feature, and that in itself can even be ambiguous (Jackson, 1988). Since there is only minimal information on manner of articulation and none concerning voicing, it is difficult to correctly identify a sound from visual information alone. Difficulty reading speech when there is no auditory signal also occurs due to characteristics of individual talkers. Other cues provided by the talker that may aid in speech perception include gestures and movements of the eyes, head and mouth. These cues are also helpful in situations involving a degraded auditory signal.

According to Grant and Seitz (1998), 'audiovisual (AV) benefit' is the amount of benefit resulting from a combination of auditory and visual cues, and this term has been used to describe the advantage of an audio-visual presentation. However, the AV benefit may depend on the relative perceptual weighting of visual and auditory cues (Hazan et al., 2006). Sumby and Pollack (1954) demonstrated that with increase in difficulty of auditory-only perception, the benefits obtained by combining the auditory and visual speech information also increased. Erber (1969) reported an improvement of 60% in word recognition scores from auditory-only condition to AV condition at -10 dB SNR for young adults. Similar results were reported using sentence materials (Middelweerd & Plomp, 1984; Summerfield, 1979). Whereas Anderson (2006) and Huffman (2007) reported that, the amount of AV integration did not vary across different auditory signal manipulations and hence, systematically removing information from the auditory stimulus does not necessarily affect the degree of integration benefit.

In the individuals with hearing impairment, AV modality has been classical approach of management in instances where auditory-only modality does not yield significant enhancement of speech perception.

1.1 Audiovisual speech perception in individuals with Hearing Impairment

The study of audio-visual integration in hearing-impaired persons is especially important, as it constitutes a case of visual input combined with reduced auditory signals. Grant and Seitz (1998) assessed integration abilities across hearing-impaired listeners using a variety of auditory-visual integration measures to establish whether integration is a process that is independent of auditory-only or visual-only processing. Congruent and discrepant nonsense syllables were degraded using a bandpass filterbank with four nonoverlapping filter bands between 300 and 6000 Hz. Congruent stimuli are described as having the auditory signal "match", or be in synchrony with, the visual articulators. Discrepant stimuli on the other hand are described as having the auditory signal and visual cue "out of sync". These stimuli can either be misaligned or have another auditory signal dubbed on to a different visual cue. These degraded syllables were then presented to listeners in the auditory (A), visual (V), and audio+visual (A+V) conditions. Results showed that even with an extremely reduced auditory signal, AV benefit was still significantly high. However, because a person's audio-only or visual-only performance could not predict their integration efficiency, Grant argued that audiovisual integration is independent of a person's ability to extract auditory and visual information from speech.

Results concerning integration independent of auditory or visual cues, however, showed little association between integration measures derived from nonsense syllable tests and those derived from sentence tests (Grant & Seitz 1998).

For all the hearing impaired individuals, and for most profoundly hearing-impaired individuals, auditory-visual (AV) speech recognition has consistently been shown to be more accurate than auditory-only (A) or visual-only (V) speech recognition. Although this is especially true when the auditory signal is distorted (e.g., due to hearing loss, environmental noise, or reverberation), the influence of visual cues on speech recognition is not limited to conditions of auditory distortion. Even with fully intact speech signals, visual cues can have an impact on recognition (Grant, Walden, & Seitz, 1998). Another example of the influence of visual speech cues on intact auditory signals occurs when listeners are asked to repeat unfamiliar phrases, as when learning a second language or when presented with grammatically complex passages.

Erber (1972) found that children with hearing impairments had increased speech perception performance on an audiovisual condition compared to an auditory alone condition, indicating that children with hearing impairments are able to use and combine both visual and auditory stimuli to process information (cited in Lachs, Pisoni, & Kirk, 2001). Lachs et al. (2001) findings were similar to Erber’s (1972) results. Lachs et al. investigated the ability of prelingually deaf children with cochlear implants to combine perceptual information from spoken language from two sensory modalities (audition and vision). The children’s performance was better on the audiovisual condition than the auditory alone and visual alone conditions (Lachs et al., 2001).

1.2 Audiovisual Speech Perception in Individuals with Auditory Neuropathy Spectrum Disorder

Auditory neuropathy spectrum disorder (ANSD) is one such disorder of the auditory system which has a cardinal feature of poor speech perception than would be expected from their degree of hearing loss (Starr, Picton, Sininger, Hood, & Berlin, 1996; Starr, Sininger & Pratt, 2000; Zeng, Oba & Starr, 2001). Majority of the studies have shown that auditory-only mode, even with sophisticated signal enhancement strategies, does not give satisfactory benefit in individuals with ANSD (Berlin, Hood, Hurely, & Wen, 1996; Hood, Wilensky, Li, & Berlin, 2004; Rance, Beer, Cone-Wesson B., Shepard, Dowell, King, Rickards, & Clark, 1999; Kumar & Jayaram, 2006; Narne & Vanaja, 2009; Vasistha & Barman, 2012)

Ramirez and Mann (2005) reported that individuals with ANSD could compensate their speech perception deficits in quiet as well as in degraded auditory conditions by focusing mainly on visual articulatory cues. The subjects in the study were divided into 3 groups; ANSD, dyslexic and control. The subjects were asked to identify the CV that was presented in the audio only and audio + video conditions. The A and AV conditions were tested at 4 different conditions (quiet, low noise, moderate noise & high noise). The identification scores of CV syllables in the visual mode acted as the baseline and the identification of CV in the other conditions was measured and compared with the baseline. Results shwed that, upon the introduction of visual cues, the performance of individuals with ANSD improved in the identification of place of articulation compared to that with auditory cues alone. Additionally, it was noted that even though individuals with ANSD showed around 50% of improvements in the perception of place cues under conditions where the auditory cues were masked, such a boost was not above the visual baseline except at the “high noise” level. This suggested that individuals with ANSD relied almost exclusively on the visual articulatory cues.

1.3 McGurk effect and Speech perception

According to the McGurk effect, if there is a mismatch or conflict between the auditory and visual information, then it disrupts the speech perception of the auditory stimulus. Mc Gurk and Mc Donald (1976) studied speech perception in adults by asking them to watch a video demonstration of a young woman’s talking head, in which repeated utterances of the syllable /ba/ had been dubbed on to lip movements for /ga/. The results showed that 98% of all adults reported the fused response, /da/.

In individuals with ANSD, speech perception in auditory mode is severely affected. Majority of the trials, stimuli are identified wrong. This is evident in the data reported by Kumar and Jayaram (2007). In most of the earlier studies, stimulus substitution in perception, similar to the above data has been accounted only to the deficits in processing temporal cues of speech. However, it is hypothesized in the present study that the stimulus confusion in perception in individuals with ANSD could be partly due to the McGurk effect. As the auditory cues perceived (that are degraded in individuals with ANSD) do not match with the visual perception, it could be leading to confusion and in turn error stimulus perception. This has not been experimentally investigated in the earlier studies.

It has been understood for many years that speech perception occurs not only for what is heard but also for what is seen (Sumby & Pollack, 1954). Speech is more intelligible when visual information is congruent with auditory information (MacDonald, Andersen & Bachmann, 2000). It has also been reported that seeing the speaker facilitates the understanding of a conversation in a noisy environment (Sumby & Pollack 1954; MacDonald et al., 2000). However, prior to 1970, it was believed that visual information aids speech perception only when auditory information is distorted. This explanation remained unchanged until McGurk and MacDonald discovered what is now simply referred to as "the McGurk effect" (McGurk & MacDonald, 1976). Audio and visual information are no longer treated as separate phenomena. Instead, speech perception is now discussed in terms of the binding of auditory and visual percepts.

The McGurk effect is a captivating experience. Laboratory studies have shown that it is also a very robust phenomenon. It occurs even when the acoustic stimulus is loud and clear, but the strength of the McGurk effect may increase when the signal-to-noise ratio (SNR) of the acoustic speech decreases (Sekiyama & Tohkura, 1991). The effect occurs even though the observer is aware of how the stimuli are constructed. Dubbing a male voice onto female articulation does not influence the strength of the McGurk effect (Green, Kuhl, Meltzoff, & Stevens, 1991), and it is not sensitive to the discrepancy in the spatial locations of auditory and visual speech (Jones & Munhall, 1997).

McGurk effect is also influenced by lexical factors. Brancazio (2004) conducted a study testing the McGurk effect on pairs of words and nonwords. The video of a /d/ sound (such as “dish”) was dubbed to the audio of a /b/ sound in the pair (such as “bish”), and the participants were required to judge whether it was a /b/ or /d sound/. It was found that when the video represented a word, McGurk effect was more likely to occur;

however when the audio was a word, the McGurk effect was less likely to occur. It was consistent with the fact that McGurk effect was a special case of visual lexical influence.

1.4 Brain Areas of Audiovisual Speech Perception

The processing of sensory signals from the outside world involves different areas in the brain. Auditory and visual information originate from two independent source of energy. Visual information originates from the transduction of photons arriving at the retina while auditory information originates from the sound pressure waves mechanically transmitted to the basilar membrane. It is argued that mutisensory integration is a highly complex mode of information processing which is not reducible to one computational stage. Neuro imaging and anatomical evidence suggest that the audiovisual integration of speech is achieved by processing in four key cortical areas, which are closely connected: posterior STS (Skipper et al., 2005) the sensory-specific auditory (Callan et al., 2004) and visual cortices (Calvert et al., 1999) and the speech motor regions (Skipper et al., 2005).

Sensory-specific cortices are said to be specialized in processing a specific type of sensory signal. The auditory cortex in the superior section of the temporal lobes specializes in processing of auditory signals (Bushara et al., 1999; Goldstein, 2007; Rauschecker & Tian, 2000; Romanski et al., 1999), whereas the occipital cortex is designated to process visual signals (Gazzaniga, Ivry, & Mangun, 2002; Goldstein, 2007; Zeki, 1978). In contrast to sensory-specific or unimodal areas, multimodal association cortices receive information from multiple senses subsequent to sensory-specific processes. Audiovisual convergence zones have been found in the temporoparietal cortex, parietal cortex, but also in premotor and prefrontal regions (Bushara et al., 1999; Calvert & Lewis, 2004; Gazzaniga et al., 2002; Romanski et al., 1999).

1.5 Auditory evoked potentials and AV Speech Perception

AV integration has been investigated using several methods, such as evoked potentials, functional magnetic resonance imaging, and intracranial local field potentials. Evoked potential data enable us to see when visual stimuli impact auditory processing. Electrophysiological (ERP) studies have found that auditory neural activity (N1 component of the ERP) induced by speech is suppressed and speeded up when a speech sound is accompanied by concordant lip movements. Hearing and seeing someone speak evokes a chain of brain responses. Once visual and auditory signals reach the ears and the eyes, these sense organs transmit their information to the sensory-specific brain areas. At some processing stage, the auditory and visual streams are then combined into a multisensory representation, as can be demonstrated by the so-called McGurk illusion (McGurk & MacDonald, 1976).

Several brain waves of the ERP are associated with the Mc Gurk effect, including N1, P2, N400, P600 and MMN (Friederici, 2002; Lebib et al., 2003; Van Wassenhove et al., 2005). The amplitude and latency of these components reflects differences in multisensory modality perception versus single modality perception as early as 150 ms following auditory stimulus onset (Van Wassenhove et al., 2005).

Seeing lip movements when listening to speech speeds, the latency of peaks as early as 10 ms post stimulation in human auditory brainstem responses (Musacchia, 2006) and from 40 to 200 ms post stimulation in cortical evoked potentials (Mottonen et al., 2002, Van Wassenhove et al., 2005). Recent evidence from cortical data suggests when subjects have one modality impaired , they use another modality effectively. Even non speech stimuli effects cortical responses to speech. Cortical activity is necessary for AV responses to occur in sub cortical structures (Musacchia, 2006).

Van Wassenhove et al., (2005) recorded EEG to the syllable /pa/, /ta/ and /ka/ from 16 subjects in an audio, visual and audiovisual paradigms using a 32-electrode system. The audio visual paradigm contained both congruent and incongruent stimuli (audio dubbed onto visual /ka/). They found that for congruent stimuli, N1 and P2 in the audiovisual condition had decreased amplitude and shortened latency compared to audio only condition. For the incongruent stimuli, N1 and P2 in the audio visual condition had the same amplitude reduction, but with the absence of the temporal facilitation.

It has been shown that audiovisual (AV) speech presentation results in a decrease in the amplitude of the N1/P2 auditory brain response (Besle et al., 2004, Van Wassenhove et al., 2005, Klucharev, Mottonen & Sams, 2003). The N1/P2 is thought to reflect the operation of early cortical mechanisms processing the initial physical attributes of an auditory stimulus (Hyde, 1997). The decrease in amplitudes does not result from a simple linear summation of N1-P2 activity with visually generated responses. Activity in the AV condition was also reduced compared with the aggregate unimodal responses. The amplitude decrease in the N1/P2 response is a reliable electrophysiological marker of the early integration of visual with auditory speech information.

However Klucharev et al., 2003 found amplitude reduction in the N1/P2 even with perceptibly mismatched AV speech. Here audiovisual pairings were of vowels that were congruent (e.g. visual /i/ with auditory /i/), or incongruent (e.g. visual /o/ with auditory /i/) with one another. Incongruent vowels did not produce a McGurk-type integrated percept. Visual information was perceived as being in clear conflict with the auditory speech signal. However, compared to the unimodal condition suppressed amplitudes in the N1/P2 were found equally for congruent and incongruent AV pairings, meaning that the effect can occur without audiovisual information producing an integrated percept. This, at least, raises the possibility that the effect may be related to presentation of information in two modalities but not necessarily to integration of AV information. Furthermore, even unimodal visual speech activates primary auditory cortex regions (MacSweeney et al., 2002).

The acoustic and visual stimuli do not have to be in exact synchrony to be integrated. The asynchrony of acoustic and visual stimuli may be as much as 240 ms (Green, 1996; Munhall et al., 1996). Furthermore, instructing subjects to respond to audiovisual stimuli based on auditory or visual information only biases responses towards the instructed modality (Massaro, 1998). However, when visual attention is directed towards a distractor stimulus presented together with the talking face, the McGurk effect is weaker indicating that visual attention modulates audiovisual speech perception (Tiippana et al., 2004).

In the domain of ERP research, the Mismatch Negativity (MMN) paradigm has come to be widely used to study speech perception. MMN has been identified as a pre-attentive ERP component (Naatanen, 2003) that is usually easily observable in an oddball paradigm, where a high ratio of similar sounds (standards) are mixed with a low ratio of deviant sounds (deviants). Colin and colleagues (Colin et al., 2002) studied the MMN in an audio, visual and audiovisual paradigm using a 6 electrode recording system (Fz, Oz, C3, C4, Ml and M2). In the audiovisual conditions, the standard and the deviant stimuli shared the same auditory information, but differed in the visual information. They reported an MMN in Fz at the P1-N1-P2 component around 50 to 170 ms in the audiovisual and audio only condition. No MMN was found in the visual only condition. The results indicated that visual information interfered with audiovisual accuracy in an early pre-attentive stage.

In another study by Kislyuk and colleagues (Kislyuk, Mottonen & Sams, 2008) observed MMN in an auditory only condition, presenting /va/ as the frequent or standard stimulus and /ba/ as the infrequent or target stimulus. In their McGurk condition, the target stimulus was the video /va/ dubbed with an auditory /ba/ (perceived as /va/); and the standard stimulus was the audiovisual /va/. Perceptually the stimuli were the same even though they had different acoustic components. The MMN observed in the Audio only condition disappeared. MMN started around 130 to 170 ms and was at various brain regions including frontal, central, occipital and central-parietal regions. In this study they used 30 electrodes and tested 11 subjects.

Saint-Amour and colleagues (2007) used high-density array. They recorded ERPs from audiovisual incongruent conditions in which the standard Audiovisual /ba/ and deviant stimulus was dubbed with audio /ba/ and visual /va/. Visual-only /ba/ and visual-only /va/ were also presented in addition to standard and deviant multisensory stimuli. The MMN was obtained by subtracting the correspondent visual-only ERP from the Audiovisual ERP. An oddball paradigm was used with deviant stimuli occurring on about 20% of the 1420 trials per modality. They found a larger left hemisphere MMN response over the frontal region that occurred at 175 to 225 ms following auditory onset, which originated from the left transverse gyrus (Broadman area 41). This was followed at 290 ms by a bilateral frontal-central activation, which was generated by a dipole in the right superior temporal gyrus (STG). After 350 ms the left hemisphere dominance (i.e., larger response over the left hemisphere) of MMN appeared again, but the dipole for wave was located in the left superior temporal gyrus. The results indicated that STG was activated during the McGurk condition, but not during the audiovisual congruent condition.

1.6 Justification for the study

According to Kumar and Jayaram (2006), prevalence of ANSD among adults in India is about 0.53% of sensory neural hearing impaired and 0.28% when all types of hearing impairments are considered. Hence, individuals with ANSD constitute an important group. In spite of understanding the disorder in terms of its speech perception difficulties, successful management of the speech perception deficits in these individuals has been an unattainable task for Audiologists. Individuals with ANSD have negligible benefit from the conventional amplification devices. Although assistive devices like FM systems have proved to be beneficial compared to conventional hearing aids, they do not address the core psychoacoustical difficulties (temporal processing deficits) encountered by these individuals and hence their utility is expected to be limited to few listening conditions only.

Individuals with ANSD are often advised to practice speech reading along with the use of listening devices (hearing aids, cochlear implants and assistive listening devices). This is done because the benefit derived from the use of the devices is minimal to moderate and often not sufficient for daily listening situation involving different background sounds. However, audiovisual benefit derived by ANSD is not well documented. Little is known about whether individuals with ANSD combine or integrate auditory and visual information in the same way as their normal hearing counterparts do.

Further, one can logically hypothesize that the stimulus confusion seen in auditory perception of individuals with ANSD may simulate McGurk effect if patients attempts simultaneous speech reading. Because the auditory cues perceived (that are degraded in individuals with ANSD) do not match with the visual perception, it could be leading to confusion and in turn error stimulus perception. This has not been experimentally investigated in the earlier studies.

The study of audiovisual perception will provide researchers with new insights into the fundamental cognitive processes used in speech processing and perception. Even normal hearing listeners benefit from speech reading cues when they are required to recognize speech in background noise or under other degraded listening conditions. Moreover, many individuals with hearing impairment and ANSD depend heavily on speech reading cues for everyday communication purposes without the availability of reliable auditory information to support speech perception. There are several electrophysiological studies of AV integration done in individuals with hearing impairment and individuals with cochlear implants. However, there is dearth of electrophysiological studies of AV integration in individuals with ANSD. Hence this study was designed.

The primary objective of the present study was to document audiovisual perception in individuals with ANSD. The purpose of the study was to verify whether the combined audio-visual perception facilitates speech perception and whether is better than perception in only visual or only auditory mode. The secondary objective was to study whether the audiovisual neural processing at the cortical level in individuals with ANSD is similar to that in normal hearing individuals.

Chapter 2

METHOD

The present study used standard group comparison research design to verify the null hypothesis that the bimodal (Audio-visual) presentation does not significantly influence speech identification and speech processing in individuals with Auditory dyssynchrony (AD). Speech identification was behaviorally assessed in audio and audio-visual conditions while speech processing was objectively studied using auditory late latency responses (ALLRs). The following method was adopted to test the hypothesis.

2.1 Participants

A total of 82 adults (48 males & 32 females) in the age range of 12 to 50 years participated in the study. They were subjected to pure-tone audiometry, immittance evaluation, speech perception in noise (SPIN) test, otoacoustic emissions (OAEs), and auditory brainstem responses (ABRs) for audiological profiling. Depending on their audiological profile, the participants were assigned into either the clinical group or the control group.

2.1.1 Clinical group

There were forty two participants in the clinical group. All of them were diagnosed as having Auditory neuropathy spectrum disorder (ANSD) by a qualified audiologist. All the participants in this group had sensorineural hearing loss of up to moderate degree with pure tone average (average of 500, 1000 & 2000 Hz) ranging from 26 to 55 dB HL. They had poor speech identification scores (Mean: 43%, Range: 0-70%) which were disproportionate to their degree of hearing loss. Identification scores were further drastically reduced (below 30% in all individuals with ANSD) in the presence of speech noise at 0dB SNR.

They did not have any history of acute or chronic ear infections (ear pain or ear discharge). The presence of external or middle ear pathology was ruled out through otoscopic examination and immittance evaluation. In the immittance evaluation, they had type 'A' tympanogram with absent ipsi- and contralateral reflexes.

All of them had normal or robust transient otoacoustic emissions (TEOAEs), but absent or abnormal ABRs, indicative of retro cochlear dysfunction. In the neurological examination, they were diagnosed as having primary auditory neuropathy by the neurologist. The Neurologist conducted a clinical examination, to evaluate the functioning different cranial nerves and to look for the signs of space occupying lesions in the brainstem . If found necessary in clinical examination, CT/MRI scan were also advised. Considering the neurological report and the audiological profile, they were diagnosed as having ANSD. The two terms, clinical group and the group with ANSD are used interchangeably in the following sections.

2.1.2 Control group

The control group included forty participants who were age and gender matched to the participants of the clinical group. In contrast to the participants in the clinical group, the participants in the control group had normal hearing sensitivity (pure-tone thresholds within 15 dB HL in the octave frequencies between 250 Hz & 8000 Hz) and normal auditory neural synchrony. Normal neural synchrony was ensured by recording replicable ABRs for clicks. They had speech identification scores of 90% or more for a phonetically balanced word list in quiet and 60% or more in the presence of speech noise at 0 dB SNR.

None of the participants had any otological symptoms (tinnitus, ear pain, ear discharge and giddiness) suggestive of conductive, cochlear or retrocochlear pathology. They had type ‘A’ tympanogram and the presence of acoustic reflexes (ipsilateral as well as contralateral) indicating normal middle ear function. They also had normal TEOAEs (SNR of greater than 6 dB SPL) indicative of normal outer hair cell functioning.

The participants in both the groups were native speakers of Kannada, a regional dravidian language spoken primarily in the state of Karnataka. They were assessed for their visual acuity using a Snellan (E) chart, to ensure normal or corrected 20/20 vision.

2.2 Stimulus generation

Natural CV syllables were used to measure beavioral speech identification as well as to record auditory late latency responses (ALLR). Voiced stops, /b/, /d/ and /g/ consonants were paired with the vowel /a/ to get CV syllables. The audio and video of the /ba/, /da/ and /ga/ utterances were used to create auditory (A), visual (V), audiovisual congruent (AVc), and audiovisual incongruent (AVi) conditions.

2.2.1 Generation of auditory stimulus

The syllables /ba/, /da/ and /ga/ were produced by a 25 yrs old female who was a native speaker of Kannada. The speaker was instructed to utter each syllable three times while maintaining a normal pronunciation, tempo, vocal intensity and neutral intonation. These stimuli were digitally recorded with a sampling frequency of 44,100 Hz and 16 bit resolution into a computer with the Adobe Audition software version 3.0. The microphone was kept at 6 inches distance from the mouth. Each syllable was approximately 250 ms in duration.

2.2.2 Generation of visual stimulus

The video of the face was recorded by a professional using digital camera, while the speaker (female) was articulating /ba/, /da/ and /ga/ utterances,. The recording was done in a room with sufficient light, at a distance of one meter. The front profile of the face and part of the upper body was covered in the video at o degree to the nase tip. This recording was digitized at 25 frames per second with a 720 x 576 graphic resolution using the Virtual dub software. Further editing was done using Adobe Audition softwaresuch that the position of the articlulators began and ended at the same neutral position for all the syllables.

2.2.3 Generation of audio-visual stimulus

Audio-visual tokens were made by dubbing the audio /ba/, /da/ and /ga/ onto video /ba/, /da/, /ga/ (congruent condition and incongruent condition) using Adobe Audition software. The total duration of the visual stimulus was 1000 ms as any legnth shorter than this only looked like a flash and it was not possible perceive the place of articulation. However, the release of the consonant was edited to occur at the auditory onset for all the visual tokens so that when the acoustic and visual stimuli were presented together, the acoustic onset occured synchronously with the release of the articulator. Because ALLR does not represent the ofset of the stimuli, it was assumed that difference in the offset of the auditory and visual stimulus does not influence the ALLR.

To generate an audio-visual congruent stimulus, each auditory stimulus was dubbed with its visual counterpart. On the other hand, to generate an audio-visual incongruent stimulus, each auditory stimulus was dubbed with a visual stimulus that is not its visual counterpart.

Table 2.1

Syllables presented in the auditory and visual modalities in different stimulus conditions used for behavioral speech identification

Auditory Only		**Visual Only**		**Audiovisual**			
				AVc		**AVi**	
A	**V**	**A**	**V**	**A**	**V**	**A**	**V**
/ba/	-	-	/ba/	/ba/	/ba/	/ba/	/da/
/da/	-	-	/da/	/da/	/da/	/ba/	/ga/
/ga/	-	-	/ga/	/ga/	/ga/	/da/	/ba/
-	-	-	-	-	-	/da/	/ga/
-	-	-	-	-	-	/ga/	/ba/
-	-	-	-	-	-	/ga/	/da/

After editing, the stimuli were randomized and presented to the participants for syllable identification (AO, VO, AVc, AVi) task. Table 2.1 gives the syllables presented in the auditory and visual modalities in different stimulus conditions used for behavioral speech identification. Table 2.2 gives the syllables presented in the auditory and visual modalities in different stimulus conditions used for the recording of ALLR.

Table 2.2

Syllables presented in the auditory and visual modalities in different stimulus conditions used for the recording of ALLR

Auditory Only	**Audiovisual**			
	AVc		**AVi**	
A	**A**	**V**	**A**	**V**
/ba/	/ba/	/ba/	/ba/	/ga/
/da/	/da/	/da/	/da/	/ba/
/ga/	/ga/	/ga/	/ga/	/da/

2.3 Test Procedure

All the audiological tests were carried out in a sound treated room. The noise level in the room was as per ANSI (1991; S3.1). The presentation of the visual stimuli was also done in an audiometric test room with good illumination.

2.3.1 Routine Audiological Evaluation

Pure-tone Audiometry: Behavioral pure-tone thresholds were obtained at octave frequencies from 250 Hz to 8 kHz through air conduction and from 250 Hz to 4 kHz through bone conduction mode. These thresholds were tracked using modified Hughson and Westlake method (Carhart & Jerger, 1959). Speech identification scores (SIS) were obtained at 40 dB SL (Re: SRT) using phonetically balanced word lists developed by Yathiraj and Vijayalakshmi (2005).

Speech Audiometry: Speech identification scores (SIS) were assessed with and without noise. The SIS in both the conditions were obtained at 40 dB above the SRT. Speech perception in noise (SPIN) scores were assessed at 0 dB SNRin the presence of speech noise. SPIN scores were used for the selection of participants. Individuals in whom SPIN score was less than 40% were considered for the study.

A calibrated two channel diagnostic audiometer (OB 922- version 2.0) coupled to its impedance matched TDH-39 head phones and B-71 bone vibrator were used to carry out pure-tone and speech audiometry.

Immittance evaluation: A calibrated immittance meter (GSI- tympstar) was used to assess the middle ear function. Each ear of the subject was tested for the type of tympanogram and the presence or absence of acoustic reflexes for a 226 Hz probe tone. Acoustic reflexes were measured at 500 Hz, 1 kHz, 2 kHz and 4 kHz pure-tones. The change of admittance by 0.03 ml was considered as presence of an acoustic reflex.

Recording of TEOAEs: Capella OAE analyzer was used to measure the TEOAEs (TEOAE amplitude, SNR and reproducibility). TEOAEs were measured using the default settings of the instrument with 260 sweeps and non-linear click trains at 75 dBpeSPL. TEOAE amplitude at each frequency band (1 kHz, 1.5 kHz, 2 kHz, 3 kHz and 4 kHz) and global TEOAE amplitude were noted. The TEOAEs were considered to be present if the overall amplitude over the noise floor was greater than or equal to 6 dB with the reproducibility of greater than or equal to 80% (Glattke, Pafitis, Cummiskey & Herrer, 1995).

Recording of ABR: Intelligent Hearing Systems (IHS) or Biologic Navigator Pro evoked potential system was used to record and analyze ABR. ER-3A insert phone was used to deliver clicks. Prior to ABR recording, electrode sites were cleaned using abrasive gel (Nuprep). Silver chloride disc electrodes were placed on the scalp at electrode placement sites with adequate amount of conductive paste. The inter electrode impedance was maintained less than 2 kOhms and intra electrode impedance was within 5 kOhm. The electrodes were taped using surgical plaster to prevent any dislodging of electrodes. The subjects were instructed to sit comfortably, close their eyes and relax on a reclining chair. They were instructed to avoid head and limb movement during testing to avoid artifacts. The ABR was recorded in every participant using the stimulus and acquisition parameters mentioned in Table 2.3.

Table 2.3

Stimulus and acquisition parameters used to record ABR

Stimulus parameter		**Acquisition parameters**	
Stimulus	Clicks	Mode of stimulation	Monaural
Polarity	Alternating	Electrode montage	Cz: +ve A1 & A2: -ve Nasion: Ground
Duration	100 µs		
Number of sweeps	1500	Filter setting	100 – 3000 Hz
Stimulus rate	11.1/sec & 90.1/sec	Transducer	Insert ear phone (ER-3A)
Intensity	90 dBnHL	Analysis window	15 ms
		Artifact rejection	>50 µV
		Gain	1,00,000
		Notch filter	On

Each recording was replicated to ensure reproducibility of ABRs. The averaged response was analyzed for peak latencies of wave I, III and V. From the peak latencies, interpeak intervals of I-III, III-V, and I-V were derived to infer the retrocochlear functional status.

2.3.2 Experiment related test procedure

The actual study related test procedure involved measurement of speech identification scores and recording of ALLR in AO, VO, AVc and AVi stimulus conditions.

Speech Identification Testing: The test stimuli were presented from a laptop computer placed at 1 meter distance from the participant. The auditory stmuli were presented at comfortable loudness levels through a high fidility Sennheisser HD449 headphones, while the visual stimuli were displayed on the computer screen at 0 degree azimuth with reference to the nose. The syllables were presented in AO, VO, AVc and AVi modalities. Each syllable was randomly trialled for 5 times in each modality and the order of modality itself was randomized across the participants. The participants were instructed to repeat the syllables heard and the responses were recorded in a scoring sheet.

The stimuli and corresponding responses were fed into a confusion matrix and analyzed using SINFA using feature info ixfer (FIX Wang & Bilger ,1973). This was done to derive place and voicing feature transmitted in different stimulus conditions in normal hearing individuals and individuals with AD.

Recording of ALLRs: Continuous EEG was acquired for the stimuli /ba/, /da/ and /ga/ in AO, AVc and AVi modalities using Neuroscan 4.4 (Compumedics, EI pasco, TX). ER-3A Insert earphones were used to deliver the auditory stimuli while 19 inches TV monitor was used to deliver the visual stimulus. All the stimuli were calibrated to 80 dBnHL and presented using Inquisit by millisecond software. The visual stimuli were triggered to the onset of the auditory stimuli in the AV stimulus conditions.

Cortical responses were acquired using Ag-AgCl electrodes (impedences <5kΩ) from five active electrodes positioned at Cz, C3, C4, Fz, Pz and Oz referenced to M1/M2. The ground electrode was postioned between Fz and FPz. A total of 300 stimuli were presented at 1.1/second rate. The different syllables used for eliciting ALLR are given in Table 2.2.

Continuous EEG was recorded with a band pass filter from 0 to 100 Hz at a sampling rate of 1000 Hz. Offline processing included dividing continuous EEG into epochs from -150 to 600 ms post acoustic onset. An artifact criterion was applied to those epochs that contain myogenic and eye blink artifacts. Any epoch with a voltage exceeding ± 50 μV was omitted from the average. The artifact-free epochs were then averaged and baseline corrected to the preauditory stimulus period.

In each participant, there were 36 averaged waves which were analyzed for marking P1, N1, P2, and N2 peaks. Each wave was independently marked by two experienced audiologists and only the markings that had consensus between the two judges were considered for data analysis. The peak latency and peak amplitude of P1, N1, P2, and N2 responses were noted.

Chapter 3

RESULTS

The results of the study are reported separately for speech identification and ALLR.

3.1 Results of Speech Identification

In the open-set identification of /ba/, /da/ and /ga/, there were total 200 presentations each, except in AVi condition. In the incongruent conditions, these syllables were presented totally 400 times in the auditory mode. Table 3.1 gives the total correct identification of each syllable summed across 40 participants of control and 40 participants of clinical group in the four stimulus conditions.

The total scores showed that, there is influence of both group as well as stimulus condition on identification of /ba/, /da/ and /ga/. Overall, identification scores were higher in control group compared to the group with ANSD. This was true in all the stimulus conditions. The difference was maximum in the AVi condition followed by AVc and AO conditions. The difference was least in VO condition.

On comparing different stimulus conditions in the control group, the scores were maximum (nearly 100%) in AO and AVc conditions but decreased in VO and AVi conditions, with least scores in the AVi condition. This was true with all three syllables. On the other hand in group with ANSD, the trend varied with the syllable. For /ba/ and /da/, score was higher in AO condition followed by AVc, VO and AVi conditions. However, for /ga/, score was maximum in AVc condition followed by VO, AO and AVi conditions.

Table 3.1

The total correct identification of syllable /ba/, /da/ and /ga/ summed across 40 participants of control and 40 participants of clinical group in the four stimulus conditions.

Group	**Stimulus**	**Stimulus Condition**			
		AO (Max=200)	**VO** (Max=200)	**AVc** (Max=200)	**AVi** (Max=400)
Control	/ba/	193	69	197	43
	/da/	189	105	193	192
	/ga/	200	60	197	100
Clinical	/ba/	48	38	50	10
	/da/	55	44	57	83
	/ga/	37	52	82	26

To statistically compare the effect of stimulus condition and group, the speech identification scores were tested on Repeated measures ANOVA (Rm ANOVA), taking group as the between-subject variable. The 2 AVi conditions (for example, auditory /ba/ with visual /da/ is one condition and that with visual /ga/ is another condition) used were treated as two separate conditions and the RmANOVA was carried out separately for the data of /ba/, /da/ and /ga/. Results of RmANOVA are given in the Table 3.2.

Table 3.2

Results of repated measures ANOVA for speech identification of /ba/, /da/ and /ga/

Stimulus	**Condition effect**		**Group effect**		**Condition X Group**
	F	**df(error)**	**F**	**df(error)**	**F**
/ba/	160.49*	4(320)	16.05*	1(80)	66.851*
/da/	99.66*	4(320)	150.15*	1(80)	40.719*
/ga/	88.09*	4(320)	38.85*	1(80)	183.461*

Note: *-p<0.001

The results of ANOVA showed significant main effect of stimulus condition as well as group on speech identification. This was true with all the three stimuli. Further, there was also a significant interaction between condition and group in all three stimuli. Therefore, RmANOVA was separately carried out for the data of the control and ANSD groups. Table 3.3 gives the results of the RmANOVA showing the effect of stimulus condition on speech identification separately in the control group and the group with ANSD.

Table 3.3

Results of repeated measures ANOVA showing the effect of stimulus condition on speech identification in the control and ANSD groups

Group	**Stimulus**	**Condition effect**	
		F	**df(error)**
Control	/ba/	191.449*	4(156)
	/da/	151.077*	4(156)
	/ga/	94.147*	4(152)
ANSD	/ba/	13.330*	4(320)
	/da/	6.358*	4(164)
	/ga/	15.804*	4(164)

Note: *-p<0.01

As there was a main effect of stimulus condition, Bonferroni test was used for pair-wise comparison of different stimulus conditions. Results of the control group can be summarized as follows

1. In the identification of /ba/, there was no significant difference between the two incongruent conditions whereas there was significant difference across all other pairs.

2. In the identification of /da/ there was no significant difference between AO and AVc condition while there was there was significant difference across all other pairs.
3. In the identification of /ga/, there was no significant difference between AO and AVc conditions. Further, there was no significant difference between Vo and the two incongruent conditions.

On the other hand, results of post hoc test in the ANSD group were slightly different and can be summarized as follows.

1. In the identification of /ba/, there was no significant difference among AO, VO and AVc conditions. In addition, there was no significant difference between the two incongruent conditions while there was significant difference between other pairs of stimulus conditions.
2. In the identification of /da/, AVi conditions were not significantly different from AO and AVc conditions. But there was a significant difference across other pairs of stimulus conditions.
3. In the identification of /ga/, AO did not differ significantly from VO and incongruent conditions. In addition, there was no significant difference between the two incongruent conditions. However, there was a significant difference across other pairs of stimulus conditions.

3.2 Results of SINFA

Sequential information transfer analysis (SINFA) derived an index representing the amount of total transmitted information in different stimulus conditions in the two groups of participants. The index represents the number of bits of information transmitted in different conditions. The index was derived for the place feature transmitted separately in the auditory and visual modalities. Table 3.4 gives the transmitted place of articulation information in 4 different stimulus conditions in the control and clinical groups.

Table 3.4

Results of SINFA showing transmitted place of articulation information in 4 different stimulus conditions in the control and clinical groups

Group	**Transmitted information**	**Stimulus Condition**			
		AO	**VO**	**AVc**	**AVi**
Control	Total	1.585	0.873	1.520	1.143
	Place-Auditory	1.585	0.835	1.510	0.019
	Place-Visual	1.585	0.835	1.510	0.382
Clinical	Total	0.222	0.699	0.880	0.678
	Place-Auditory	0.131	0.650	0.805	0.319
	Place-Visual	0.131	0.650	0.805	0.012

The data in Table 3.4 shows that, over all, the transmission index (TI) of the place of articulation feature was better in the control group than in group with ANSD . In the control group, TI was maximum in AO and AVc conditions followed by VO condition. TI was least in AVi condition. On the contrary, in the ANSD group, TI of place of articulation feature was maximum in AVc condition followed by VO and AO conditions. Similar to control group, TI in group with ANSD was least in AVi condition. The maximum difference between the two groups was evident in AO condition.

3.3 Results of ALLR

ALLR were recorded from 6 electrode sites using the stimuli given in Table 2.2. The peak latency and peak to peak amplitudes were noted down from the ALLRrecorded from Cz only. The responses from the other electrode sites were used only to generate topographic maps and were not tested statistically. The results of the ALLR are separately reported for the latency and amplitudes.

3.3.1 Results of Peak Latency

Table 3.5 gives the mean and standard deviation (SD) of peak latencies of P1, N1, P2 and N2 elicited by syllables /ba/, /da/, /ga/ as auditory stimuli in AO, AVc and AVi stimulus conditions in control and ANSD groups. The inspection of the mean data reveals that,

1) 1) When elicited by /da/ and /ga/, the latencies were earlier in the AVi condition than in the AO and AVc conditions. Whereas no such trend was seen for /ba/. The mean latencies were prolonged in group with ANSD compared to control group. This was true for all four waves (P1, N1, P2 & N2) elicited by all the three stimuli (/ba/, /da/ & /ga/) in all 3 conditions i. e., AO, AVc and AVi.

To verify whether the observed differences in the mean latencies were statistically significant, RmANOVA was carried out with stimulus condition as repeating variable and group as between-subject variable. Results of ANOVA are given in Table 3.6.

Table 3.5

Mean and standard deviation (SD) of peak latencies(ms) of P1, N1, P2 and N2 elicited by syllables /ba/, /da/, /ga/ as auditory stimuli in AO, AVc and AVi stimulus conditions in control group and group with ANSD

Wave	**Condition**	**/ba/**				**/da/**				**/ga/**			
		Control		**ANSD**		**Control**		**ANSD**		**Control**		**ANSD**	
		Mean (ms)	**SD**	**Mean (ms)**	**SD**	**Mean (ms)**	**SD**	**Mean (ms)**	**SD**	**Mean (ms)**	**SD**	**Mean (ms)**	**SD**
P1	Auditory	74.89	26.26	99.12	29.98	71.66	24.66	107.31	50.23	64.85	18.21	84.44	31.31
	Congruent	72.34	25.08	106.18	36.74	76.87	102.56	102.56	18.06	73.92	24.30	87.83	21.68
	Incongruent	77.48	21.61	107.75	28.01	76.66	25.86	92.50	25.94	56.39	20.58	64.33	23.89
N1	Auditory	140.90	23.42	155.46	47.37	142.03	19.80	189.44	39.72	127.43	24.87	158.75	27.60
	Congruent	140.78	33.18	141.65	69.66	151.92	23.21	171.00	20.23	141.34	22.62	176.65	29.03
	Incongruent	135.24	52.49	152.96	69.59	146.34	27.79	162.88	23.73	117.12	30.50	129.65	19.29
P2	Auditory	206.00	22.22	233.00	38.57	211.57	30.43	263.66	47.23	205.21	28.30	234.57	28.60
	Congruent	210.57	32.46	235.00	41.00	218.03	25.27	243.55	33.66	215.75	26.48	253.31	32.61
	Incongruent	227.03	28.60	254.50	25.69	203.42	38.17	236.83	25.29	186.03	47.20	207.78	27.54
N2	Auditory	297.07	24.13	305.33	38.57	294.00	34.50	317.88	43.43	303.84	58.95	308.89	30.35
	Congruent	294.57	35.28	315.77	31.29	308.76	26.87	318.11	40.76	299.09	38.36	336.26	48.02
	Incongruent	315.82	37.16	324.55	34.74	292.92	34.69	302.00	24.20	275.12	51.43	287.78	50.73

Table 3.6

Results of repeated measures ANOVA for peak latencies of P1, N1, P2 and N2 waves

Wave	Stimulus	Condition effect		Group effect	
		F	df(error)	F	df(error)
P1	/ba/	0.509	2(86)	28.756**	1(43)
	/da/	0.468	2(76)	20.181**	1(38)
	/ga/	11.496**	2(88)	8.208**	1(44)
N1	/ba/	0.278	2(114)	2.065	1(57)
	/da/	2.526	2(84)	24.778**	1(42)
	/ga/	27.75**	2(100)	25.424**	1(30)
P2	/ba/	5.844**	2(88)	23.122**	1(44)
	/da/	3.102	2(84)	33.88**	1(42)
	/ga/	19.489**	2(98)	19.906**	1(49)
N2	/ba/	5.844*	2(88)	3.963	1(44)
	/da/	5.14*	2(84)	7.379*	1(42)
	/ga/	7.993**	2(98)	4.218*	1(49)

Note:-p<0.05 and **-p<0.01*

Results of ANOVA showed significant main effect of condition on peak latencies of all four waves (P1, N1, P2 & N2) when elicited by /ga/. However, the main effect of condition was present only for N2 when elicited by /da/ and on P2 as well as N2 when elicited by /ba/.

Because there was a main effect of condition, pair-wise differences were tested on Bonferroni test. Results of this can be summarized as follows.

1) There was a significant difference between AO and AVc only in N1 and P2 latencies elicited by /ga/.

2) There was a significant difference between AO and AVi in latencies of all waves showing the main effect of condition except N2 elicited by /ga/.
3) There was a significant difference between AVc and AVi conditions in latencies of all waves showing the main effect of condition except N2 elicited by /ba/.

On the other hand, there was a significant difference between the two groups in peak latencies of all the waves, in all three conditions for all three stimuli, except N1 and N2 for /ba/.

3.3.2 Results of Peak Amplitude

Table 3.7 gives the mean (µV) and standard deviation of peak to peak amplitudes of P1-N1, N1-P2, and P2-N2 elicited by syllables /ba/, /da/, /ga/ as auditory stimuli in AO, AVc and AVi stimulus conditions. The data of both control group and group with ANSD are given in the table.

Inspection of the data in Table 3.5 shows that the amplitudes are higher in control group compared to group with ANSD in all three stimulus conditions when elicited by /ba/. But there is no such trend observed for /da/ and /ga/ stimuli. Further, there was no noticeable trend observed in the way amplitudes varied among the three stimulus conditions.

Repeated measures ANOVA was carried out to test the significance of differences in the mean peak amplitudes across conditions and between the two groups. Results of ANOVA are shown in Table 3.8. Results showed that there was no significant effect of either stimulus condition or group on the peak to peak amplitudes of ALLR. This was true for all the three stimuli.

Table 3.7

Mean and standard deviation (SD-in paranthesis) of peak to peak amplitudes of P1-N1, N1-P2, and P2-N2 elicited by syllables /ba/, /da/, /ga/ as auditory stimuli in AO, AVc and AVi stimulus conditions in control group and group with ANSD

Parameter	**Condition**	**/ba/**		**/da/**		**/ga/**	
		Control	**ANSD**	**Control**	**ANSD**	**Control**	**ANSD**
		Mean (μV)	**Mean (μV)**	**Mean (μV)**	**Mean (μV)**	**Mean (μV)**	**Mean (μV)**
P1-N1	Auditory	4.15 (2.05)	3.07 (1.53)	3.47 (1.32)	3.18 (1.55)	3.46 (1.51)	3.24 (1.40)
	Congruent	4.19 (2.03)	3.68 (1.74)	3.73 (1.71)	3.36 (1.34)	3.31 (1.36)	3.49 (1.42)
	Incongruent	3.77 (2.55)	3.43 (1.30)	3.75 (1.60)	3.30 (2.11)	3.43 (1.74)	3.57 (2.22)
N1-P2	Auditory	4.09 (2.77)	3.92 (2.46)	2.75 (1.57)	3.24 (1.95)	2.79 (2.20)	2.89 (1.65)
	Congruent	4.61 (1.58)	3.76 (3.56)	2.58 (1.98)	3.08 (1.77)	3.09 (2.21)	3.26 (1.47)
	Incongruent	3.88 (2.74)	3.26 (1.50)	2.11 (2.23)	3.50 (1.94)	2.65 (1.74)	3.47 (1.82)
P2-N2	Auditory	2.63 (1.94)	1.65 (1.55)	2.28 (1.76)	2.33 (1.95)	2.65 (2.00)	2.09 (1.44)
	Congruent	3.11 (2.32)	2.72 (4.16)	1.76 (1.63)	2.07 (1.46)	1.97 (1.82)	1.90 (1.77)
	Incongruent	2.30 (2.84)	1.69 (1.62)	1.42 (1.51)	2.20 (1.21)	2.14 (2.02)	2.32 (1.77)

Table 3.8

Results of repeated measures ANOVA for peak to peak amplitudes of P1-N1, N1-P2, and P2-N2 complexes

Parameter	**Stimulus**	**Condition effect**		**Group effect**	
		F	**df(error)**	**F**	**df(error)**
P1-N1	**/ba/**	0.462	2(96)	2.914	1(48)
	/da/	0.395	2(84)	0.918	1(42)
	/ga/	0.171	2(100)	0.007	1(50)
N1-P2	**/ba/**	0.953	2(96)	0.859	1(48)
	/da/	0.155	2(86)	3.870	1(43)
	/ga/	0.879	2(100)	0.613	1(50)
P2-N2	**/ba/**	2.286	2(96)	1.620	1(48)
	/da/	1.642	2(82)	1.048	1(41)
	/ga/	1.104	2(96)	0.126	1(48)

Note: $p>0.05$ in all the comparisons

3.3.2 Results of Scalp Topography

The topographic distributions of the grand averaged responses for all the ALLR peaks were analysed. Figure 3.1 to 3.3 shows the topographic maps generated from the responses elicited for /ba/, /da/ and /ga/ respectively.

The topographic maps for stimuli /ba/ Figure 3.1 . The findings are as follows.

1. The scalp distributions in the control group for the ALLR peaks were similar for the AO and AVc conditions for all the peaks except for P1, which suggested a

wider activation in the AVc condition. The findings in the AVi condition are suggestive of more activation in the left temporo-parietal regions for the N1 peak.

2. The group with ANSD showed a wider activation for the P1 latency region in the AVc condition compared to the AO condition. The results are suggestive ofa greater left temporal activation in the AO condition and posterior left temporo-parietal activation in the AVc condition for the N1 peak. The AVi condition showed a generalised activation over the mid coronal and frontal regions for the N1.
3. The control group and the group with ANSD when compared differed in the activation maps for the N1 peak. The control groups showed great activation in the left temporal region along with activation over the whole scalp. Whereas, the group with ANSD showed a more generalised activation. This was true for all the conditions. The difference in activation across conditions followed similar trend in both the groups

The topographic maps for stimuli /da/ Figure 3.2. The findings are as follows.

1. The AO condition in the control groups suggest greater activations in temporal regions along with mid-coronal and mid-frontal regions. The AVc condition showed a more left temporal and left posterior parietal activation for the N1 peak. The AVi condition showed similar activation as the AVc condition, with slightly greater activation in the right mid-coronal regions. Further, topographic maps indicate greater activation in the left posterior temporo-parietal regions for the P2 and N2 peaks in the AVi condition compared to the AO and AVc condition.

2. The group with ANSD showed a slightly greater activation in the right temporal regions with generalised mid-coronal activation in the AO condition for N1 peak, while in the AVc condition there was a left temporal activation. The P2 peak showed a greater frontal activation in AO condition and showed a left temporal activation in the AVc condition. The AVi condition created greater activation in both the temporal regions for the N1 peak, and for the P2, activation was similar to the AO condition.

The control group and the group with ANSD differed in the activation maps for the N1 peak. The maps derived from the control groups suggest great activation in the left temporal region along with activation over the whole scalp. Whereas, the group with ANSD suggest a more generalised activation. In addition, the difference in activation across conditions was more prominent in the group with ANSD than the control group.

The topographic maps for stimuli /ga/ (Figure 3.3). The activation map for N2 peak in the AVc condition has not been included as the N2 peak was not clear in the grand average waveform.The findings are as follows.

1. In the AO condition, more generalised activation for the N1 , with a slightly greater activation in the left temporal region, was observed. The AVcondition showed greater activation in the left temporo-parietal regions for the N1, and P2 than the AO condition. The AVi condition was more similar to the AO condition for the N1 peak and similar to AVc for the P2 and N2 peaks with slightly greater left temporal activation.
2. The group with ANSD showed a more generalised activation in mid coronal and frontal regions in the AVc condition compared to a greater frontal activation with

generalised mid-coronal activation in the AVc condition for the N1 peak. The AVi condition suggested a left temporal activation in for the P1 as opposed to the frontal activation in the AO and AVc condition. The N1 in AVi condition indicate a wider activation compared to AVc condition. The P2 in the AVi condition suggest a greater frontal activation compared to the AO and AVC condition.

3. Comparison of the control group and group with ANSD showed that there was a greater left temporal activation in control group compared to the group with ANSD. Further, the AVi condition in the group with ANSD suggest a wider activation compared to the other conditions, whereas, the control group indicate a wider activation in the AVc condition compared to the other conditions.

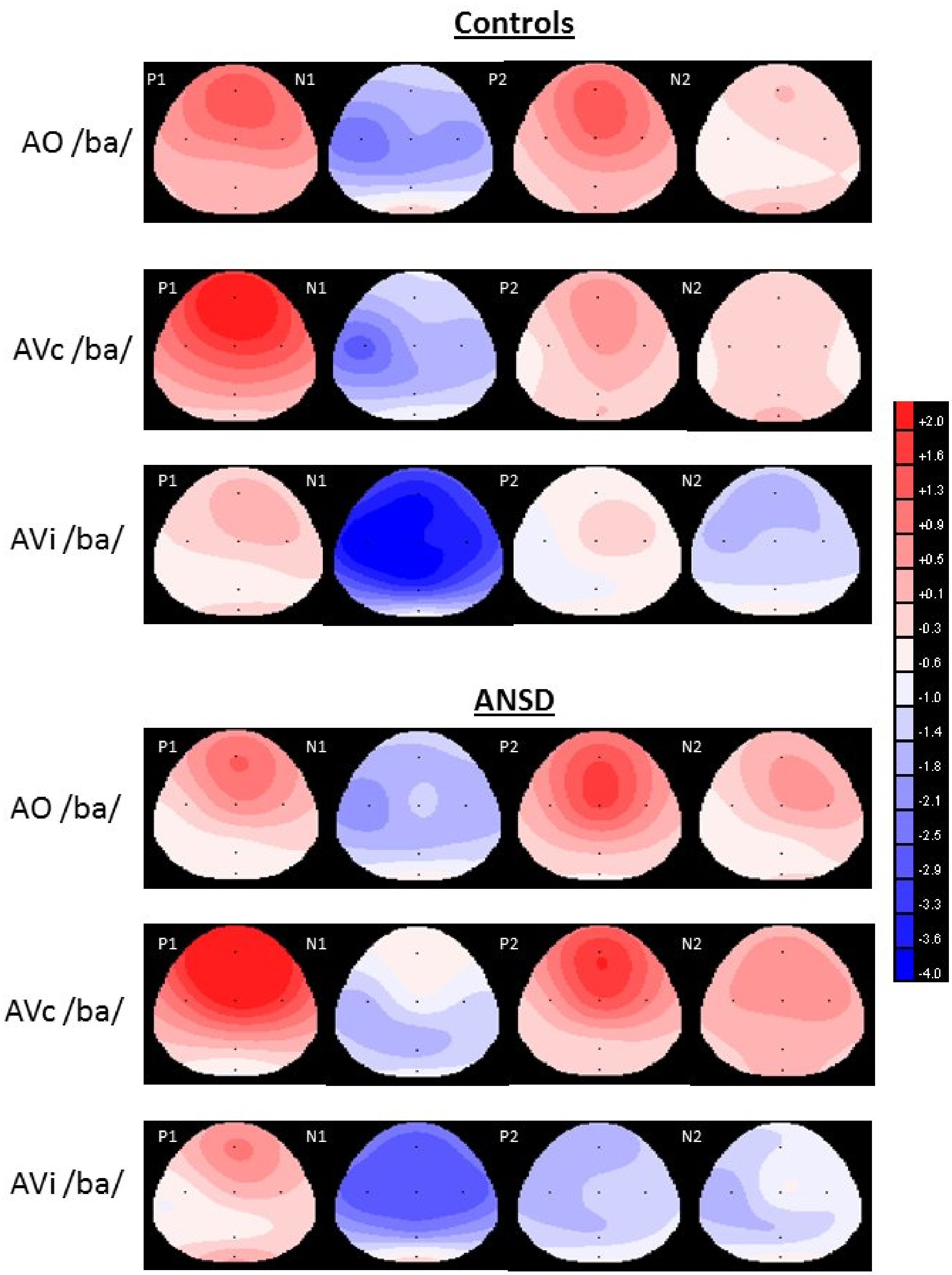

Figure 3.1

Scalp topographic maps generated for the ALLRs recorded in three stimulus conditions in the two groups for stimulus /ba/.

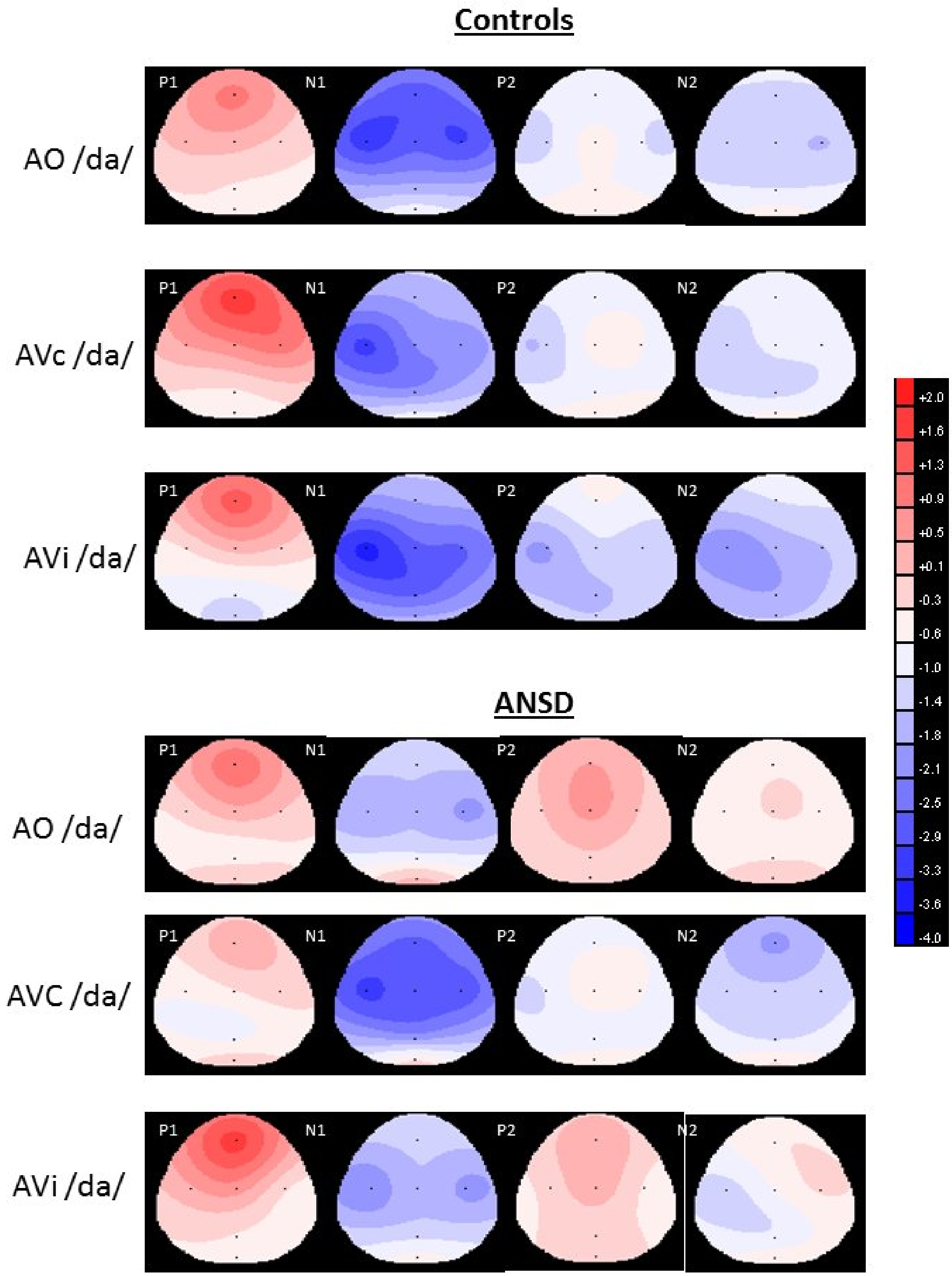

Figure 3.2

Scalp topographic maps generated for the ALLRs recorded in three stimulus conditions in the two groups for stimulus /da/.

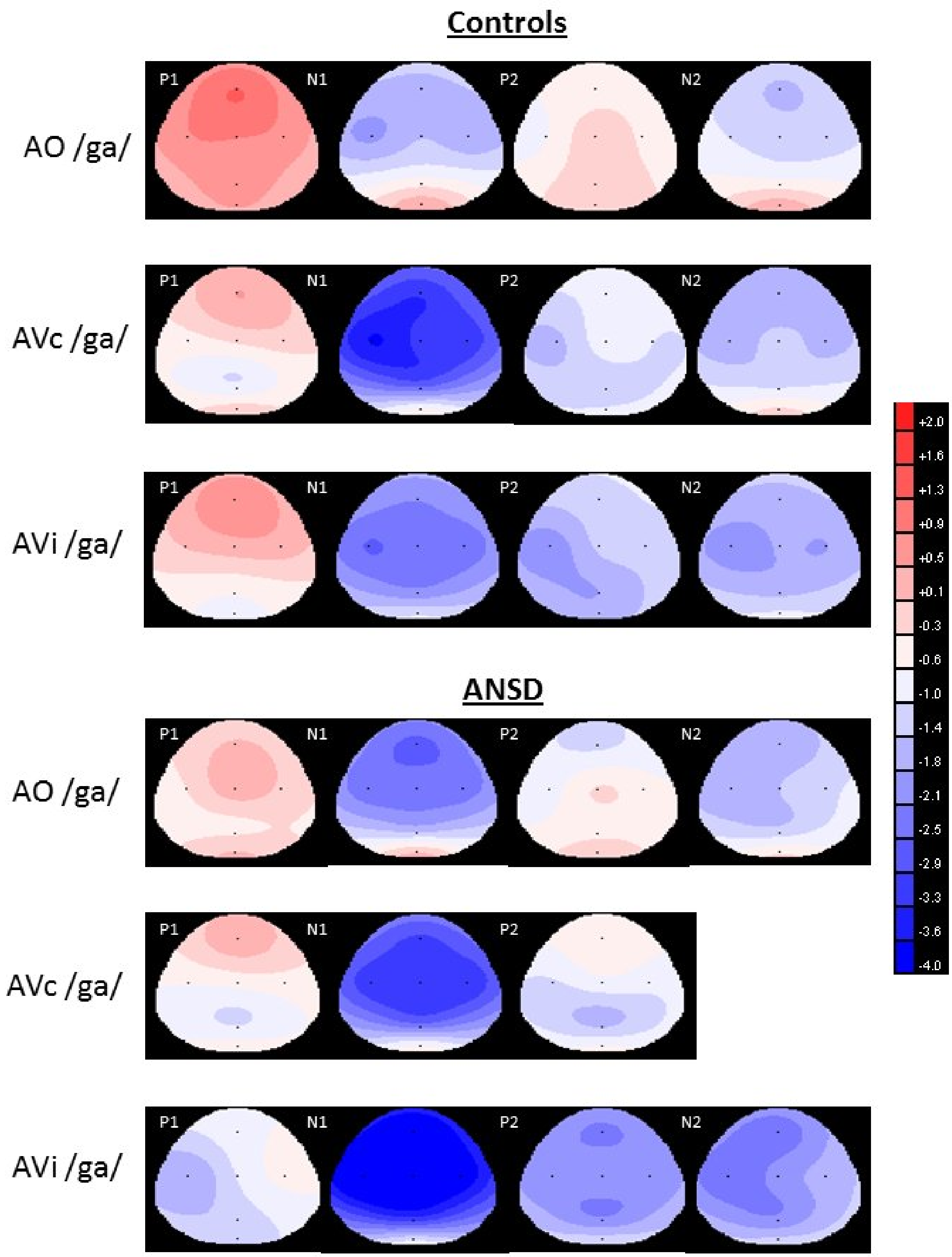

Figure 3.3

Scalp topographic maps generated for the ALLRs recorded in three stimulus conditions in the two groups for stimulus /ga/.

Chapter 4

DISCUSSION

The results of the present study are discussed under this chapter in light of Audio-visual (AV) modality for the facilitation of speech perception in individuals with Auditory neuropathy spectrum disorder (ANSD). It was of interest in the present study to investigate whether individuals with ANSD utilize and integrate cues from both auditory and visual modality or are they dependent on the dominant modality. Further it was necessary to experimentally document whether cues from the two modalities interfere with each other and in turn reduce speech perception, due to mismatch that could be created by ANSD. The findings of the study are discussed under the following headings.

1. Effect of modality on speech identification in individuals with ANSD
2. Effect of modality on features perceived in individuals with ANSD
3. Effect of modality on Auditory late latency responses (ALLR)

4.1 Effect of Modality on Speech Identification in Individuals with ANSD

Speech identification was documented in four stimulus conditions; auditory-only (AO), visual-only (VO), audiovisual congruent (AVc) and audiovisual incongruent (AVi). Irrespective of the condition, individuals with ANSD showed significantly poorer speech identification compared to normal hearing individuals. The poor speech perception in the auditory modality is a cardinal feature of ANSD and the finding is in agreement with the earlier studies (Starr et al., 1996; Starr, Sininger & Pratt, 2000; Zeng, Oba & Starr, 2001; Kumar & Jayaram, 2006; Narne & Vanaja, 2009; Vasistha & Barman, 2012). The auditory perceptual deficits in subjects with ANSD are reported to be mainly due to the disruption of temporal cues (Kraus et al., 2000; Starr et al., 1991). Kumar and Jayaram (2005) reported that the individuals with ANSD have severely affected temporal processing. Rance, McKay and Grayden (2004) have also reported that individuals with ANSD perform poorly for the task involving timing cues and they found a correlation between the temporal processing abnormalities and speech perception abilities.

In addition to the above, previously reported speech perception difficulties in the auditory modality, the findings of the present study show that individuals with ANSD lag behind normal hearing individuals, not just in the auditory mode but also invisual-only (VO), audiovisual congruent (AVc) and audiovisual incongruent (AVi) conditions. In the identification of all three syllables (/ba/, /da/ & /ga/) AO did not differ from VO. That means, speech perception through auditory modality and visual modality was comparable in individuals with ANSD. Whereas in normal hearing individuals, auditory modality was way better compared to visual modality. Comparison of speech identification of the two groups in VO condition, showed that speech identification was significantly poorer in group with ANSD compared to normal hearing group. This was an interesting result. Ideally, one would expect visual speech perception scores to be comparable between the two groups considering the ANSD is disorder in the auditory domain. However, in actual, the findings showed that even visual processing is affected in individuals with ANSD.

The three speech sounds taken in the study differed only in terms of place of articulation. Therefore, if the AO and VO scores did not differ in individuals with ANSD, it means that the amount of place of articulation cues conveyed in auditory mode and visual mode is similar in individuals with ANSD. From these results one can infer that both auditory and visual processing were affected in individuals with ANSD. Therefore, one needs to assess speech perception through visual modality also, to ensure the benefit that could be derived, before advising speech reading in these individuals.

Ramirez and Mann (2005) reported that ANSD could compensate their speech perception deficits in quiet as well as in degraded auditory conditions by focusing mainly on visual articulatory cues. Upon the introduction of visual cues, individuals with ANSD show definite improvement in identifying place of articulation compared to their performance with auditory cues alone. However, based on the present findings one can infer that speech perception will not be enhanced to any great extent if focused on visual articulatory cues.

The inference is also supported by the finding that AVc did not facilitate speech perception compared to AO condition. Considering that in AVc condition, the place of articulation cues are simultaneously provided through both auditory and visual modalities, one would expect that AVc gives better scores than AO condition. However, the improvement was seen only in the identification of /ga/. This means that the visual modality provided additional cues and facilitated speech perception only for /ga/ while for the other two consonants, the place information provided by the two modalities were similar. On the other hand, in the normal hearing group, speech perception in AVc and AO conditions were similar due to ceiling effect.

Based on the findings of AVc condition, one can infer that audio visual perception does not significantly improve the perception of speech identification compared to only the auditory modality. However, one must realize that the speech identification testing in the present study was conducted in audiometric rooms where the signal to noise ratio is quite high as the rooms are sound treated. The same findings may not hold correct in the adverse listening conditions where the auditory cues in individuals with ANSD are further compromised. Considering that visual cues provide same amount of place of articulation information, audiovisual speech perception should still be the choice of management in individuals with ANSD. However, the audiologists should be aware and must counsel that clients that benefit derived would be more in the adverse listening conditions.

In the incongruent conditions, both normal hearing individuals and individuals with ANSD performed poorer in the speech identification. The individuals with ANSD although performed poorer than the normal group. The reduced performance of the group with ANSD in AVc and AVi conditions can be interpreted in two ways. One, the cues provided by the auditory and visual modality are lesser in individuals with ANSD and therefore the poorer performance in the audiovisual conditions. Secondly, it may be due to their inability to integrate the information from the two modalities.

Overall from the speech identification performance, one can infer that the audiovisual speech perception may be a choice of management in individuals with ANSD. However, it is not advantageous over auditory modality in the quiet environments.

4.2 Effect of Modality on Features Perceived in Individuals with ANSD

The performance on speech identification task could be misleading while understanding effects of modality on place of articulation feature. Therefore, the Sequential information transfer analysis (SINFA) was carried out to derive the place of articulation feature provided different stimulus conditions and in the two groups. Feature analysis showed a picture different from what was evident through speech identification performance.

Similar to speech identification performance, SINFA showed that group with ANSD perceived place feature to much lesser extent than the normal group. The place of articulation feature is primarily cued by F2 transition. The individuals with ANSD because of the temporal processing deficits secondary to auditory dys-synhrony are not able to process and perceive the difference in the F2 transitions of /ba/, /da/ and /ga/. The results of the study are in agreement with the earlier reports of by Kumar and Jayaram (2007).

The place of articulation feature was maximally distorted in AVi condition, both in normal and ANSD groups. This decrement can be explained through McGurk phenomena. According to McGurk effect, if there is a mismatch or conflict between the auditory and visual information, then it disrupts the speech perception of the auditory stimulus. Mc Gurk and Mc Donald (1976) studied speech perception in adults by asking them to watch a video demonstration of a young woman's talking head, in which repeated utterances of the syllable /ba/ had been dubbed on to lip movements for /ga/. Results showed that 98% of all adults reported the fused response, /da/. McGurk phenomena is primarily demonstrated for place of articulation feature.

The feature analysis showed that the place of articulation feature was perceived poorer in the auditory mode compare to VO and AVc conditions. Although reduced index in VO condition in the group with ANSD compared to normal group, points towards associated visual processing deficits, visual modality provided better place information than the auditory modality in group with ANSD . Further in AVc condition, the perception of the place feature was facilitated compared to AO and VO conditions. Thus, one can infer that the individuals with ANSD are able to integrate the information from the two modalities and utilize it for the perception of place of articulation feature.

The differences in the results of speech identification and feature analysis could be because of the voicing feature. That is, of one identifies /ba/ as /pa/, he/she is scored wrong on an identification performance while is scored correct in terms of identification of place of articulation feature. Therefore, feature analysis is better approach to study the effects of modality on speech perception.

Overall from the findings of SINFA, it can be inferred that AV speech perception facilitates the identification of place of articulation of consonants. However, this may not assure correct identification of the consonant as a whole. For the perception of place of articulation, AV speech perception is helpful management strategy in individuals with ANSD.

4.3 Effect of Modality on Auditory Late Latency Responses (ALLR) in individuals with ANSD

Among the forty two individuals with ANSD, only thirty three of them had recordable ALLR. Earlier report by Narne & Vanaja (2008) have shown that individuals with ANSD can have absent or poor ALLR and has been attributed to the dys-synchronous firing of the cortical neurons). The results of ALLR showed prolonged latencies in group with ANSD compared to normal hearing individuals, irrespective of the stimulus and the stimulus condition. The prolonged latencies either indicate slower processing or dys-synchronous firing individuals with ANSD. ALLR are obligatory potentials that represent the registration of the stimulus at the primary and secondary auditory cortex. Therefore, the absent or the prolonged ALLRs indicate poorer representation of the speech in the cortical regions in individuals with ANSD.

Results of latencies also showed the effect of modality. Latencies were earlier in AVi condition compared to the other conditions. Although the exact reason for this is not clear, one can speculate that the irregularities created by the incongruency trigger earlier latency responses in ALLR. The findings also showed that the congruent condition of /ga/ elicited prolonged ALLR compared to AO condition. This is in contradiction to the earlier study by Van Wassenhove, Grant and Poeppel (2005) who reported earlier latencies and smaller amplitudes in AVc condition.

The amplitudes of ALLR were similar between the two groups and across three stimulus conditions. This could be either due to higher standard deviation or because physiologically there are no differences in the amplitudes.

Scalp topography indicated differences between the two groups and across stimulus conditions. Scalp topography overall indicated lower and diffuse neural activity in individuals with ANSD. This could be because of the de-synchronous firing of the cortical neurons in group with ANSD and there may not be specialized areas in them as in normal hearing individuals.

Chapter 5

SUMMARY AND CONCLUSIONS

There has been a considerable increase in the prevelance of Auditory neuropathy spectrum disorders (ANSD). It is a known fact that the benefit of listening devices in Individuals with ANSD is very limited and these individuals are often advised to practice speech reading along with the use of listening devices. However, little is known about whether they combine or integrate auditory and visual information in the normal way or not.

Hence, the primary objective of the present study was to verify whether the combined audio-visual perception facilitates speech perception or not?. The secondary objective was to study whether the audiovisual neural processing at the cortical level in individuals with ANSD is similar to that in normal listeners.

The study included 82 adults (48 males & 32 females) in the age range of 12 to 50 years. Of the 82 participants 42 (clinical group) were individuals with ANSD and 40 (control group) were normal hearing individuals. Audio and video recording of CV syllables /ba/, /da/ and /ga/ were done to create auditory (AO), visual (VO), audiovisual congruent (AVc), and audiovisual incongruent (AVi) conditions. Behavioral speech identification was done and auditory late latency responses (ALLR) were recorded in all the four conditions.

Analysis of the behavioral speech identification data revealed that the group with ANSD performed poorer than the control group in all the stimulus conditions for all the syllables. The maximum difference between the two groups was evident in AO condition.

Speech perception in AO and VO conditions was similar in individuals with ANSD. Further, speech identification in VO was significantly poorer in group with ANSD compared to normal hearing group. This implies that even visual processing is affected in individuals with ANSD. In the normal hearing group, speech perception in AVc and AO conditions were similar due to ceiling effect. However, in group with ANSD, this did not happen, which reveals that addition of visual information did not facilitate speech perception.

AVi conditions yielded poorest performance in both the groups. The individuals with ANSD performed poorer than the normal group. The reduced performance of the ANSD group in the audiovisual conditions could be due to the auditory and visual cues are distorted and due to their inability to integrate the information from the two modalities.

Sequential information transfer analysis (SINFA) was performed to derive the place of articulation feature provided in different stimulus conditions and in the two groups. Results of this revealed that the ANSD group perceived place feature to much lesser extent than the normal group. The place of articulation feature was perceived poorer in AO compared to VO and AVc conditions suggesting visual modality providing better place information than the auditory modality in group with ANSD .

In AVc condition, the perception of the place feature was facilitated compared to AO and VO conditions, suggesting that the individuals with ANSD are able to integrate the information from the two modalities to some extent. The differences in the results of speech identification and feature analysis could be because of the voicing feature. The place of articulation feature was maximally distorted in AVi condition, both in normal group and group with ANSD which could be attributed to McGurk effect.

The ALLR waves were analyzed for the peak latency, amplitude and scalp topography. The results showed that the mean latencies were prolonged in group with ANSD compared to control group the indicating poorer representation of the speech in the cortical regions in individuals with ANSD. Latencies were earlier in AVi condition compared to the other conditions. This might be because of the irregularities created by the incongruency trigger earlier latency responses in ALLR. The findings also showed that the congruent condition of /ga/ elicited prolonged ALLR compared to AO condition.

Results of comparison of amplitudes revealed that there was no significant effect of either stimulus condition or group on the peak to peak amplitudes of ALLR. This was true for all the three stimuli. Scalp topography overall indicated lower and diffuse neural activity in individuals with ANSD which could be because of the de-synchronous firing of the cortical neurons in group with ANSD.

To conclude, the study shows scientific evidence facilitation of speech perception particularly place of articulation of consonants, in the audio-visual modality. Therefore, for the perception of place of articulation, AV speech perception can be a helpful as a management strategy in individuals with ANSD. This finding rejects the apriori assumption that individuals with ANSD could have inherent McGurk effect due perceptual deficits in the auditory mode.

Another important finding is that the visual processing is also declined in individuals with ANSD. Therfore, although deficient, auditory modality plays a role in the perception of speech sounds and therby facilitating it with suitable amplification device is necessary in individuals with ANSD. Also, because of the deficient visual processing, one may expect lesser facilitation even in the bimodal perception than what is expected if the visual processing was to be completely normal.The information derived in this research can be utilized in predicting the prognosis with AV mode and counseling the clients accordingly.

REFERENCES

American National Standards Institute. (1991). Maximum permissible ambient noise levels for audiometric test rooms. ANSI S3.1-1991, New York: American National Standards Institute.

Benoît, C., Mohamadi, T., & Kandel, S. (1994). Audio-Visual Intelligibility of French speech in noise. *Journal of Speech & Hearing Research*, *37*, 1195-1203.

Berlin, C. I., Hood, L, J., Hurley, A., & Wen, H. (1994). Contralateral suppression of otoacoustic emissions: An index of the function of the medial olivocochlear system. *Otolaryngology Head Neck Surgery, 100*, 3-21.

Besle, J., Fort, A., Delpuech, C., & Giard, M. H. (2004). Bimodal speech: Early suppressive visual effects in human auditory cortex. *European Journal of Neuroscience*, *20,* 2225-2234.

Brancazio, L. (2004). Lexical influences in audiovisual speech perception. *Journal of Experimental Psychology: Human Perception and Performance, 30*, 445-463.

Bushara K. O., Weeks, R. A., Ishii, K., Catalan, M. J., Tian, B., Rauschecker, J. P., Hallett, M. (1999) Modality-specific frontal and parietal areas for auditory and visual spatial localization in humans. *Nature Neurosciences*, *2*, 759-766.

Callan, D. E., Jones, J. A., Munhall, K., Kroos, C., Callan, A. M., Vatikiotis-Bateson, E. (2004). Multisensory integration sites identified by perception of spatial wavelet filtered visual speech gesture information. *Journal of Cognitive Neuroscience, 16*, 805–816.

Calvert, G. A., Brammer, M.J., Bullmore, E.T., Campbell, R., Iversen, S. D., & David, A. S. (1999). Response amplification in sensory-specific cortices during crossmodal binding. *Neuroreport, 10,* 2619-2623.

Calvert, G. A., Lewis, J. W. (2004). Hemodynamic studies of audio-visual interactions. In G. A. Calvert, C. Spence, B. Stein, (Eds.) *Handbook of multisensory processing* (pp 483–502). Cambridge, MA: MIT Press.

Carhart, R., & jerger, J. F. (1959). Preferred method for clinical determination of pure tone thresholds. *Journal of Speech and Hearing Disorders, 24, 330-345.*

Cienkowski, K. M., & Carney, A. E. (2002). Auditory-visual speech perception and aging. *Ear and Hearing, 23*(5), 439-449.

Colin C, Radeau M, Soquet A, Colin F, Deltenre P. (2002). Mismatch negativity evoked by the McGurk-MacDonald effect : Evidence for a phonological representation within the auditory sensory short term memory. *Clinical Neurophysiology, 113,* 495-506.

Erber, N. P. (1969). Interaction of audition and vision in the recognition of oral speech stimuli. *Journal of Speech and Hearing Research,12,* 423-425.

Erber, N. P. (1972). Auditory, visual, and auditory-visual recognition of consonants by children with normal and impaired hearing. *Journal of Speech and Hearing Research, 15,* 413–422.

Frederici, A. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences, 6,* 78-84.

Grant, K. W., Walden, B. E., & Seitz, P. F. (1998). Auditory-visual speech recognition by hearing-impaired subjects: Consonant recognition, sentence recognition, and auditory-visual integration. *Journal of the Acoustical Society of America*, *103*, 2677–2690.

Grant, K.W., & Seitz, P.F. (1998). Measures of auditory-visual integration in nonsense syllables and sentences. *The Journal of the Acoustical Society of America*, 104, 2438-2450.

Grant, K.W., & Seitz, P.F. (2000). The use of visible speech cues for improving auditory detection of spoken sentences. Journal of Acoustical Society of America, (108), 1197–1208.

Green, K. P. (1998). The use of auditory and visual information during phonetic processing: Implications for theories of speech perception. In R. Campbell, B.

Dodd, & D. Burnham (Eds.), *Hearing by eye II: Advances in the psychology of speechreading and auditory–visual speech* (pp. 3–26). Hove, England: Psychology Press.

Green, K. P., Kuhl, P. K., Meltzoff, A. N., & Stevens, E. B. (1991). Integrating speech information across talkers, gender, and sensory modality: Female faces and male voices in the McGurk effect. *Perception & Psychophysics*, *50*, 524-536.

Gzzaniga, M. S., Ivry, R. B., & Mangun, G. R. (2002). *Cognitive Neuroscience: The biology of the mind* (2nd ed.). New York: W.W.Norton.

Hazan, V., Sennema, A., Faulkner, A., Ortega-Llebaria, M., & Chung, H. (2006). The use of visual cues in the perception of non-native consonant contrasts. *Journal of the Acoustical Society of America, 119* (3), 1740 – 1751.

Hood, L. J., Wilensky, D., Li, L., & Berlin, C. I. (2004). The role of FM technology in the management of patients with auditory neuropathy/dys-synchrony. *Proceedings of the International Conference on FM Technology*, Chicago, Illinois.

Huffman, C. (2007). The Role of Auditory Information in Audiovisual Speech Integration. *Senior Honors Thesis, the Ohio State University*.

Hyde, M. (1997). The N1 response and its applications. *Audiology Neurootology*, *2*(5), 281-307.

Jackson, P. L. (1988). The theoretical minimal unit for visual speech perception: Visemes and coarticulation. *The Volta Review, 90* (5), 99-114.

Jones, J. A., & Munhall, K. G. (1997) The effects of separating auditory and visual sources on audiovisual integration of speech. *Canadian Acoustics, 25*(4), 13-19.

Kislyuk D. S., Möttönen, R., & Sams, M. (2008) Visual processing affects the neural basis of auditory discrimination. *Journal of Cognitive Neuroscience. 20*, 2175–2184.

Klucharev, V., Mottonen, R., & Sams, M. (2003). Electrophysiological indicators of phonetic and non-phonetic multisensory interactions during audiovisual speech perception. *Cognitive Brain Research*, *18,* 65-75.

Kumar, A. U., & Jayaram, M. (2006). Prevalence and audiological characteristics in individuals with auditory neuropathy/dys-synchrony. *International Journal of Audiology, 45*, 360-366.

Kumar, A. U., & Jayaram, M. (2007.). Perception of some temporal parameters of speech in individuals with auditory dys-synchrony. *Unpublished dissertation submitted to the University of Mysore*, Mysore.

Lachs, L., Pisoni, D. B., Kirk, K. I. (2001). Use of audiovisual information in speech perception by prelingually deaf children with cochlear implants: A first report. *Ear and Hearing, 22,* 236–251.

Ladefoged, P. (2006). A *Course in Phonetics-Fifth Edition.* Boston : Wadsworth.

Lebib, R., Papo, D., de Bode, S., & Baudonniere, P. M. (2003). Evidence of a visual-to-auditory cross-modal sensory gating phenomenon as reflected by the human p50 event-related brain potential modulation. *Neuroscience Letters, 341,* 185–188.

MacDonald, J., & McGurk, H. (1978). Visual influences on speech perception processes. Perception and Psychophysics, *24*(3), 253-257.

MacDonald, J., Andersen, S., & Bachmann, T. (2000). Hearing by eye: How much spatial degradation can be tolerated? *Perception*, *29,* 1155-1168.

MacSweeney, M., Calvert, G.A., Campbell, R., McGuire, P.K., David, A.S., Williams, S.C., et al., (2002). Speechreading circuits in people born deaf. *Neuropsychologia, 40* (7), 801- 807.

Masaro, D. W, & Cohen, M. M. (2000). Test of audiovisual integration efficiency within the framework of the fuzzy logical model of perception. *Journal of the Acoustical Society of America*, *108,* 784-789.

Massaro, D. W., & Stork, D. G. (1998). Speech recognition and sensory integration. *American Scientist, 86,* 236–244.

McGurk ,H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature, 264*, 746-8.

Middleweerd, M., & Plomp, R. (1984). The effect of speechreading on the speech reception threshold of sentences in noise. *Journal of the Acoustical Society of America, 82,* 2145-2147.

Mottonen, R., Krause, C. M., Tiippana, K., & Sams, M. (2002) Processing of changes in visual speech in the human auditory cortex. *Cognitive Brain Research, 13,* 417-425.

Musacchia, G, Sams, M, Nicol, T., & Kraus, N. (2006). Seeing speech affects acoustic information processing in the human brainstem. *Experimental Brain Research, 168,* 1-10.

Näätänen, R. (2003). Mismatch negativity: Clinical research and possible applications. *International Journal of Psychophysiology, 48,* 179-188.

Narne, V. K., & Vanaja, C. S. (2009). Perception of envelope enhanced speech in the presence of noise by individuals with auditory neuropathy. *Ear and Hearing, 30*, 136-142.

Ramirez, J., & Mann, V. A. (2005). Using auditory-visual integration to probe the basis of noise-impaired speech perception in reading disability and auditory neuropathy *Journal of the Acoustical Society of America*, *118*, 122-133.

Rance, 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*(3), 20(3):238-52.

Rauschecker, J. P., Tian, B. (2000) Mechanisms and streams for processing of "what" and "where" in auditory cortex. *Proceeding of the National Academy of Sciences, U S A 97,*11800-11806.

Romanski, L. M., Tian, B., Fritz, J., Mishkin, M., Goldman-Rakic, P. S., Rauschecker, J. P. (1999) Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. *Nature Neuroscience, 2,* 1131-1136.

Rosenblum, L. D. (2005). Primacy of multimodal speech perception. In B. Pisoni & R. Perez (Eds.), *Handbook of Speech Perception* (pp. 51-78). MA, USA: Blackwell Publishing.

Rosenblum, L.D., Schmuckler, M. A., & Johnson, J. A. (1997). The McGurk effect in infants *Perception & Psychophysics, 59*(3), 347-35.

Saint-Amour, D., De Sanctis, P,, Molholm, S., Ritter, W., & Foxe, J. J. (2007). Seeing voices: high-density electrical mapping and source-analysis of the multisensory mismatch negativity evoked during the McGurk illusion. *Neuropsychologia, 45,* 587–597.

Sekiyama, K., & Tohkura, Y. (1991). McGurk effect in non-English listeners: Few visual effects for Japanese subjects hearing Japanese syllables of high auditory intelligibility. *Journal of the Acoustical Society of America*, *90,* 1797-1805.

Skipper, J. I., Nusbaum, H. C., Small, S. L. (2005). Listening to talking faces: motor cortical activation during speech perception. *Neuroimage, 25*, 76-89.

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

Starr, A., Sininger, Y., & Pratt, H. (2000). The varieties of auditory neuropathy. *Journal of Basic Clinical Physiology and Pharmacology, 11*(3), 215-230.

Sumby, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *Journal of the Acoustical Society of America, 26*(2), 212-215,

Summerfield, Q.A. (1979). Use of visual information for phonetic perception. *Phonetica,36,* 314-331.

Tiipana, K., Andersen, T. S., & Sams, M. (2004). Visual attention modulates audiovisual speech perception . *European Journal of Cognitive Psychology*, *16*, 457-472.

van Wassenhove, V., Grant, K. W., & Poeppel, D. (2005). Visual speech speeds up the neural processing of auditory speech. *Proceedings of the National Academy of Sciences, U.S.A., 102*, 1181–1186.

Vasistha, S., & Barman, A. (2012). Perception of spectrally enhanced speech through companding in individuals with auditory dys-synchrony. *An unpublished dissertation submitted to the university of Mysore, Mysore.*

Wang, M. D., & Bilger, R. C. (1973). Consonant confusions in noise: A study of perceptual features. *Journal of Acoustical Society of America, 54*, 1248-1266.

Zeki, S. M., (1978). Functional specialization in the visual cortex of the rhesus monkey. *Nature*, *274,* 423-428.

Zeng, F. G., Oba, S., & Starr, A. (2001). Suprathreshold processing deficits due to desynchronous neural activities in Auditory Neuropathy. In D. J. Breebaart, A. J. M. Houstma, A. Kohlrausch, Prijs, V.F., & Schoonhoven, R. (Eds.) *Physiological and Psychophysical Bases of Auditory Function*, (pp. 365-372), Maastricht, Netherlands: Shaker publishing BV.