

Speech perception in individuals with auditory dys-synchrony

U A KUMAR¹, M JAYARAM²

¹Department of audiology and speech language pathology, Kasturba Medical College (a unit of Manipal University), Mangalore, and ²Department of Speech Language Pathology and Audiology, National Institute of Mental Health and Neuroscience, Bangalore, India

Abstract

Objective: This study aimed to evaluate the effect of lengthening the transition duration of selected speech segments upon the perception of those segments in individuals with auditory dys-synchrony.

Methods: Thirty individuals with auditory dys-synchrony participated in the study, along with 30 age-matched normal hearing listeners. Eight consonant–vowel syllables were used as auditory stimuli. Two experiments were conducted. Experiment one measured the ‘just noticeable difference’ time: the smallest prolongation of the speech sound transition duration which was noticeable by the subject. In experiment two, speech sounds were modified by lengthening the transition duration by multiples of the just noticeable difference time, and subjects’ speech identification scores for the modified speech sounds were assessed.

Results: Subjects with auditory dys-synchrony demonstrated poor processing of temporal auditory information. Lengthening of speech sound transition duration improved these subjects’ perception of both the placement and voicing features of the speech syllables used.

Conclusion: These results suggest that innovative speech processing strategies which enhance temporal cues may benefit individuals with auditory dys-synchrony.

Key words: Auditory Disorder; Auditory Brainstem Responses; Speech Perception; Auditory dys-synchrony; Auditory Neuropathy

Introduction

Auditory dys-synchrony, also known as auditory neuropathy, is a recently described hearing disorder which has unique physiological and perceptual consequences. A major characteristic of this disorder is the disrupted activity of the auditory nerve in the presence of normal or near-normal cochlear outer hair cells.¹ Auditory brainstem responses (ABRs) are either absent or severely abnormal in these patients. Advances in the accurate assessment of outer hair cell function have made it easier to diagnose auditory dys-synchrony. Persons with this condition generally show poor speech perception² and derive only limited benefit from hearing aids.³ In the present study, speech perception was studied in individuals with auditory dys-synchrony by lengthening the transition duration and the voice onset time. The effects of lengthened transition duration upon speech perception are presented here. The effect of voice onset time modification, and of combined modification of voice onset time and transition duration, and its relationship with temporal processing ability are reported in a companion paper.

In individuals with sensorineural hearing loss, the prevalence of auditory dys-synchrony has been estimated at approximately 0.54 per cent.² Prevalence rates are higher (around 10 per cent) in hyperbilirubinaemic infants.⁴ The exact lesion site and pathophysiology of auditory dys-synchrony are not yet completely understood.

The aetiological factors affecting auditory dys-synchrony are only now being established, and appear diverse. Neonates exposed to hyperbilirubinaemia and anoxia seem to be at increased risk of auditory dys-synchrony.⁴ Genetic factors have also been identified; Starr *et al.*⁵ reported a novel mutation in the MPZ gene in a family with hereditary motor sensory neuropathy and deafness.

The mechanism of auditory dys-synchrony may include: loss of inner hair cells;⁶ dysfunction of synaptic junctions between inner hair cells and the auditory nerve;⁷ and auditory nerve demyelination or axonal loss, or both.⁸

Individuals with auditory dys-synchrony show marked deficits in processing temporal (i.e. time-based)

auditory information, but have relatively good processing of auditory intensity and frequency information.⁹ It is now well established that the speech identification problems of individuals with auditory dys-synchrony are disproportionate to their degree of hearing loss.¹ In fact, it can be said that poor speech perception which is disproportionate to the pure tone hearing thresholds is the cardinal characteristic of auditory dys-synchrony. In affected individuals, the degree of speech perception appears to depend on the extent of suprathreshold temporal distortion of speech cues, rather than on access to the speech spectrum.^{9–13} This is in contrast to subjects with cochlear hearing loss, who typically demonstrate loudness recruitment, broadening of auditory filters and normal processing of temporal information, at least at high sensation levels. Persons with cochlear hearing loss derive significant benefit from hearing aids which employ nonlinear compression circuits. All such hearing aids assume abnormal functioning of outer hair cells. Hence, these aids are seldom of benefit for individuals with auditory dys-synchrony, who have normal outer hair cell functioning.¹⁴ These hearing aids do not change the temporal parameters of speech sound (using linear amplification) or reduce fluctuations when a non-linear amplitude-compression circuit is employed.¹⁵ Other appliances and strategies that are helpful to individuals with auditory dys-synchrony include frequency modulation systems, cochlear implants, perceptual training, speech reading and cued speech. Cochlear implantation may be a viable option for some patients with auditory dys-synchrony.^{15–19} However, many patients, particularly in developing countries, will find the cost of a cochlear implant prohibitive.

Previous studies on the speech perception and psychophysical abilities of individuals with auditory dys-synchrony have suggested that temporal processing is markedly affected. In fact, there is reasonable evidence to suggest that individuals with auditory dys-synchrony find it more difficult to perceive short duration dynamic sounds than long duration steady sounds.^{11,13}

Therefore, the purpose of the present study was to investigate the effect of the transition duration of speech segments (a short duration temporal cue) upon the speech perception of individuals with auditory dys-synchrony. The specific aims of the study were (1) to measure 'just noticeable differences' in speech segment transition duration, using speech stop consonants, in individuals with auditory dys-synchrony compared with normal hearing persons, and (2) to investigate the effect of lengthened speech segment transition durations upon the speech perception of individuals with auditory dys-synchrony.

Materials and methods

Experiment one: subjects

Two groups of subjects participated in experiment one. The first group consisted of 30 individuals with auditory

dys-synchrony (19 men and 11 women; age range 16–30 years, mean age 22.4 years). The second group consisted of 30 age- and gender-matched normal hearing individuals. All auditory dys-synchrony subjects were recruited from the audiology department of the All India Institute of Speech and Hearing, Mysore.

Prior to recruitment, all subjects underwent a structured interview conducted by the first author, an audiologist and speech-language pathologist. These interviews established that none of the subjects had any history of middle-ear disease, noise exposure, ototoxic drug usage, or a previous history of sound amplification or hearing rehabilitation.

Of the 30 auditory dys-synchrony subjects, 25 reported onset of auditory dys-synchrony between the ages of 16 and 20 years, three reported onset at 22–24 years, and the remaining two reported onset at 25 years. No specific aetiology could be identified in most subjects. Two participants reported that the problem had started after they had given birth to their first child. All the subjects had clinically normal speech.

The results of various audiological tests conducted on the auditory dys-synchrony subjects are given in Table I.

Figure 1 shows representative ABR and otoacoustic emission waveforms for a subject with auditory dys-synchrony.

The conditions and procedures used for audiological evaluation were as described by Kumar and Jayaram.² The methods used to measure contralateral suppression of otoacoustic emissions were the same as described by Kumar and Vanaja.²³

Before selection for the study, all auditory dys-synchrony subjects underwent an ENT examination to exclude any external or middle-ear pathology. Similarly, all subjects underwent a neurological examination, conducted by a qualified neurologist, to exclude any peripheral neuropathy or space-occupying lesion.

Subjects in the normally hearing group had hearing thresholds within 15 dB HL at octave frequencies between 250 and 8 kHz, and had normal results on immittance evaluation.

All subjects were native speakers of Kannada, a Dravidian language spoken by approximately 55 million people in South India.²⁴

Experiment one: stimuli preparation

Consonant–vowel syllables with voiceless stop consonants (i.e. velar /ka/, alveolar /ta/, retroflex /ʈa/ and bilabial /pa/) and their voiced cognates were used in the study. These syllables, spoken in isolation by a 25-year-old, male, native Kannada speaker, were digitally recorded onto a data acquisition system at 16 kHz sampling frequency with a 16-bit analogue-to-digital converter.

The recorded syllables were then altered using the pitch synchronised overlap and add technique, accessed via the Praat software package (Amsterdam, The Netherlands).²⁵ The pitch synchronised overlap and add technique enables temporal lengthening of

TABLE I
AUDIOLOGICAL PROFILE OF SUBJECTS WITH AUDITORY DYS-SYNCHRONY

S no	Age (y)/sex	PTA (RE/LE; dB HL)	SIS (RE/LE; %)	TEOAE (RE/LE; dB SNR)	ABR	Acoustic reflex	OAE supprn (dB)	Audiogram
1	16/F	43/43	45/45	22.3/24.5	A	A	0.2	Peaked
2	16/M	13/16	84/84	22.4/21.9	A	A	0.0	Peaked
3	30/M	23/35	80/80	21.9/22.3	A	A	0.4	Rising
4	24/F	31/46	38/36	26.5/22.3	A	A	0.0	Rising
5	16/M	55/63	0/0	22.5/25.8	A	A	0.1	Peaked
6	26/M	40/51	4/4	22.7/26.3	A	A	0.0	Peaked
7	23/F	41/26	5/5	29.9/20.5	A	A	0.0	Peaked
8	23/M	76/73	0/0	16.5/17.6	A	A	0.1	Peaked
9	27/M	31/31	80/80	16.7/17.5	A	A	0.0	Rising
10	23/M	23/23	86/84	20.6/29.1	A	A	0.3	Peaked
11	24/M	53/53	8/8	27.8/28.9	A	A	0.0	Peaked
12	25/F	43/43	68/68	11.5/20.7	A	A	0.0	Peaked
13	28/M	31/31	80/80	12.5/13.9	A	A	0.3	Rising
14	25/F	36/31	0/0	23.5/22.9	A	A	0.2	Peaked
15	16/M	36/31	60/60	21.9/29	A	A	0.2	Peaked
16	18/M	31/31	76/80	22.7/26.5	A	A	0.3	Peaked
17	22/M	50/53	0/0	24.6/25.9	A	A	0.0	Rising
18	24/F	3/3	90/90	24.7/25.9	A	A	0.0	Peaked
19	23/M	15/13	95/95	27.9/25.4	A	A	0.2	Peaked
20	26/M	56/65	0/0	16.7/17.6	A	A	0.1	Peaked
21	20/F	48/51	0/0	15.7/13.4	A	A	0.2	Rising
22	18/F	43/43	26/40	12.3/15.9	A	A	0.3	Peaked
23	16/F	13/16	84/84	15.8/17.9	A	A	0.1	Peaked
24	18/M	25/25	82/82	23.4/27.8	A	A	0.1	Rising
25	19/M	31/46	38/36	28.4/21.1	A	A	0.0	Rising
26	30/M	52/63	0/0	22.2/25.6	A	A	0.0	Peaked
27	28/M	40/51	4/4	24.6/24.7	A	A	0.0	Peaked
28	27/F	31/26	6/6	22.5/23.5	A	A	0.0	Peaked
29	22/F	76/26	0/0	22.4/23.8	A	A	0.0	Peaked
30	20/M	31/31	62/80	20.9/19.7	A	A	0.1	Rising

S no = subject number; y = years; PTA = pure tone average (of 0.5, 1 and 2 kHz); RE = right ear; LE = left ear; SIS = speech identification score (for monosyllables);²² TEOAE = transient evoked otoacoustic emissions; SNR = signal to noise ratio; ABR = auditory brainstem response; OAE supprn = otoacoustic emission suppression; F = female; M = male; A = absent; peaked = sharp peak at single frequency with worsening of thresholds at immediately adjacent frequencies by more than 10 dB; rising = decrease in threshold by 5 dB or more per octave

stimulus in the time domain while still preserving most of the physical characteristics of the stimulus, such as spectral shape, amplitude distribution and periodicity.²⁶ It was difficult to accurately segment natural speech. Identification of transition was confirmed as follows. It was noted that, in transition regions, the sound frequency changed by more than 200 Hz (in 25–30 ms depending on speech sound). However, in the steady-state portions of the vowel, the change in sound frequency was less than 40 Hz in 30 ms. Furthermore, pitch synchronised overlap and add technique processing was applied within 40 ms of the stimulus onset. Thus, we were reasonably certain that only the format transition of the syllables had been lengthened. It is highly unlikely that any significant part of the steady-state segment of the vowel was also lengthened, although such a possibility cannot be excluded.

Figure 2 shows the waveform, spectrogram and spectrum of the unmodified syllable /ba/, while Figure 3 shows the same parameters for the modified syllable /ba/.

Experiment one: procedure

Both the unmodified and modified speech material was presented, in turn, through a loud speaker placed at an

angle of 0° azimuth and a distance of 1 m away from the subject. The presentation level of the sound signals was kept constant at 40 dB SL (the reference was average thresholds at 500 Hz, 1 kHz and 2 kHz). If 40 dB SL could not be reached due to audiometric limits or subject discomfort, testing was performed at an intensity 10 dB below the level of discomfort. This was required in three participants (subjects number 8, 20 and 29).

The loud speaker output was calibrated at the beginning of the experiment and then regularly during the test period, using a sound level meter (Quest 1800; Quest technologies, Wisconsin, USA) and microphone (Quest 4180).

Subjects were tested individually in a sound treated room. Sound signals were fed through a personal computer, at a sampling frequency of 44 kHz, to an audiometer (Maico MA-53; Minnesota, USA). Subjects received the sound signals through a loudspeaker connected to the audiometer. The loudspeaker was positioned as described above. The 'just noticeable difference' was determined using an adaptive tracking technique (Parameter estimation through sequential testing (PEST)) with an AX same-difference discrimination paradigm (where A = anchor stimulus and X = variable stimulus): subjects were asked to indicate

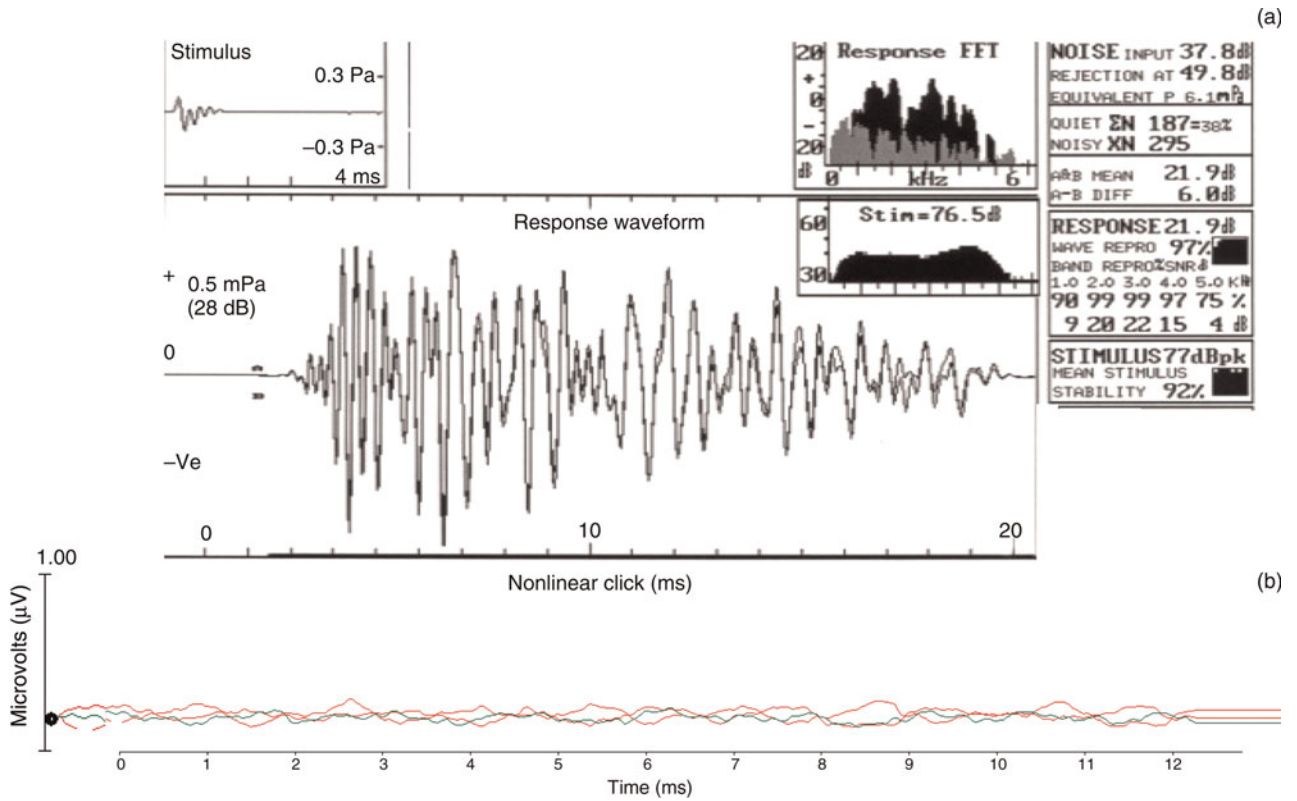


FIG. 1

Representative (a) transient evoked otoacoustic emission and (b) auditory brainstem response waveforms for a subject with auditory dys-synchrony. FFT = Fast Fourier transformation; stim = Stimulus

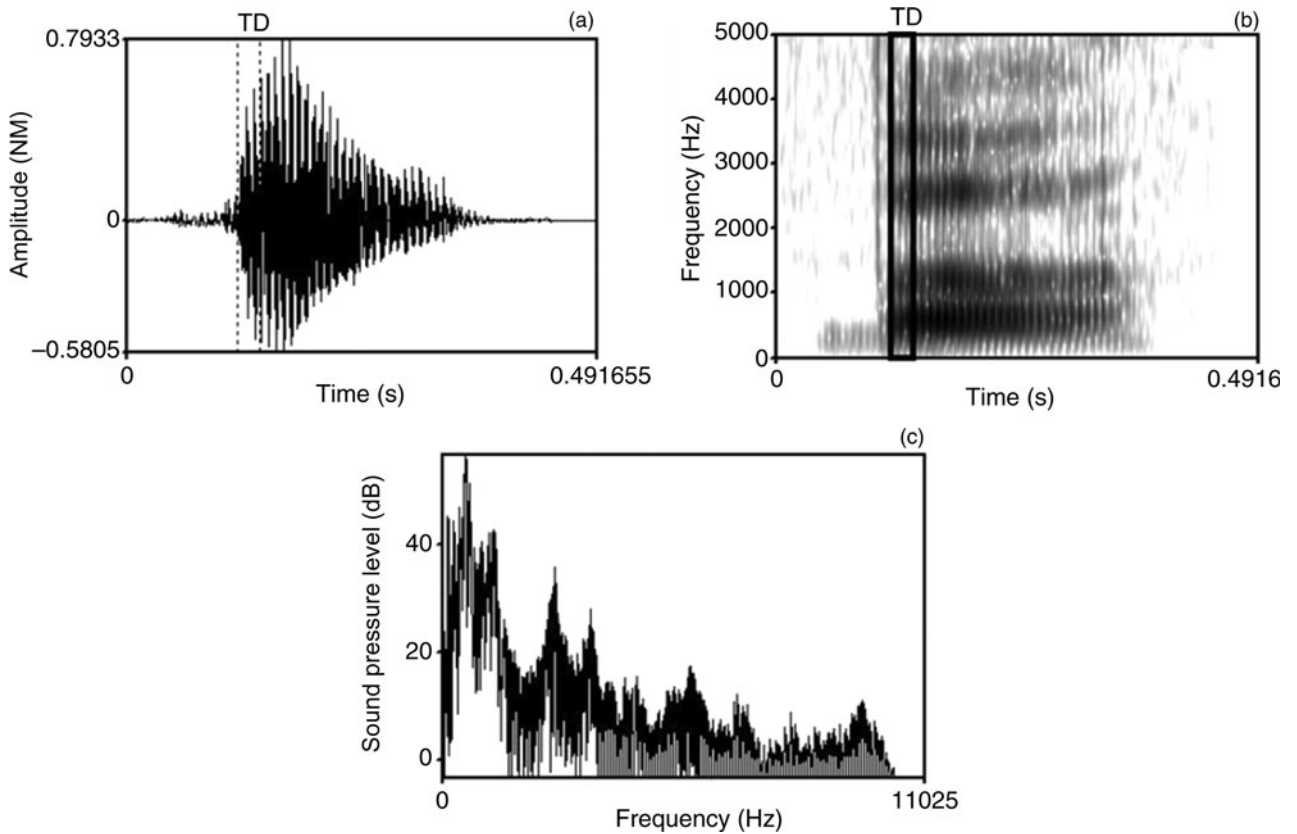


FIG. 2

(a) Spectrogram of unmodified speech stimulus /ba/. (b) Waveform of unmodified speech stimulus /ba/; the transition duration (TD) is marked with a rectangle. (c) Frequency spectrum of the unmodified speech stimulus /ba/

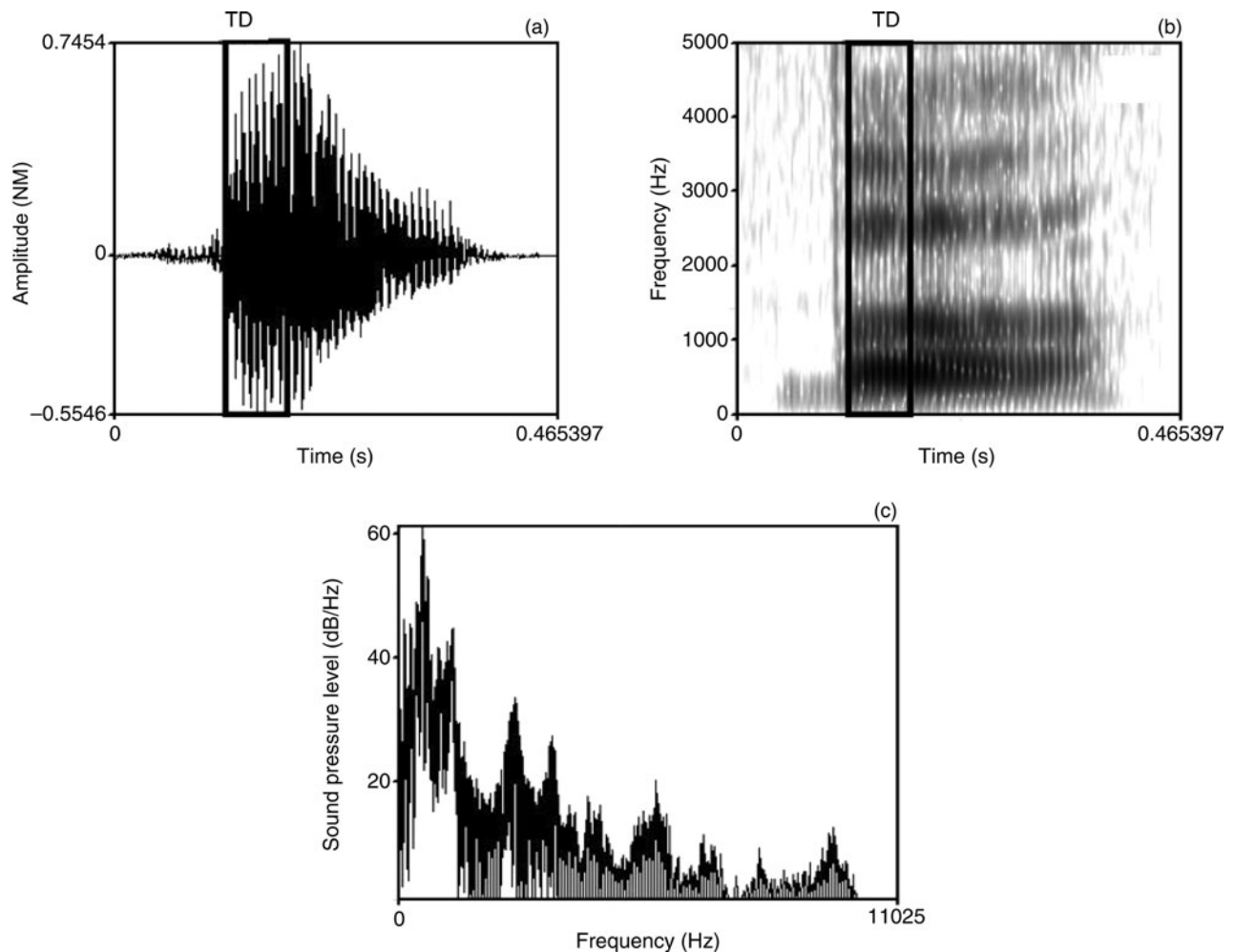


FIG. 3

(a) Spectrogram of modified speech stimulus /ba/ (transition duration (TD) lengthened). (b) Waveform of modified speech stimulus /ba/; TD is marked with a rectangle. (c) Frequency spectrum of the modified stimulus /ba/.

whether A was the same as X or not. The inter-stimulus interval between the anchor and variable stimuli was 500 ms. The step size and the direction of the variable stimulus were changed as determined by “parameter estimation through sequential testing” technique.²⁷ Each subject’s just noticeable difference time was determined by calculating the difference in transition duration between the anchor stimulus and the variable stimulus which was required to achieve a 69 per cent correct response. Test trials included an equal number of catch trials; these consisted of either two identical anchor or two identical non-anchor stimuli.

Experiment two: speech identification

Speech identification scores for unmodified speech stimuli were obtained for both the auditory dys-synchrony subjects and the normal hearing subjects. Ten repetitions of each of the eight unmodified speech signals were randomly presented and speech identification scores were noted. Stimuli were played through a personal computer at a sampling frequency of 44.1 kHz, and then fed into a calibrated clinical audiometer (Maico MA-53). Subjects received the

stimuli through a loudspeaker connected to the audiometer. The loudspeaker was positioned at an angle of 90° azimuth and a distance of 1 m from the subject. The sound presentation level was kept constant at 40 dB SL for all subjects (the reference was average thresholds at 500 Hz, 1 kHz and 2 kHz).

Written responses were obtained from subjects if they were literate. If subjects were illiterate, the experimenter and another native Kannada speaker noted down the subject’s responses independently; only those responses for which there was 100 per cent agreement between the two observers were used for the further analysis.

Subsequently, the speech signal transition duration was lengthened by a time period equal to multiples of the just noticeable difference time, and the effect of this lengthening upon the subject’s speech identification score was determined. Only subjects with auditory dys-synchrony participated in this experiment. Four modified speech sounds were generated to evaluate the effect of lengthening of the transition duration upon the subject’s perception of consonant–vowel syllables. The first modified speech sound was the syllable

/pa/ with the transition duration lengthened by a time period equal to one just noticeable difference (the mean just noticeable difference of the normal listeners was used). Similarly, three more modified speech sounds were generated wherein the transition duration was lengthened by time periods equal to two, three or four just noticeable difference times.

Each modified speech sound was presented to the subject 10 times. Presentation stopped when the subject identified the correct sound, or the sound which elicited the closest response to the target sound as determined by feature analysis. Feature analysis considered voicing and placement of articulation information. The order of presentation of sound objects was randomised to minimise the practice effect.

Results

Experiment one

Repeated measures analysis of variance revealed a statistically significant difference between the just noticeable difference times for altered syllable stop consonant transition durations, comparing the auditory dys-synchrony and normal hearing groups ($F(1, 63) = 4471; p < 0.01$).

Independent sample *t*-testing was conducted to establish the statistical significance of differences between the just noticeable difference times for each speech sound signal, comparing the two groups. Results revealed a statistically significant difference

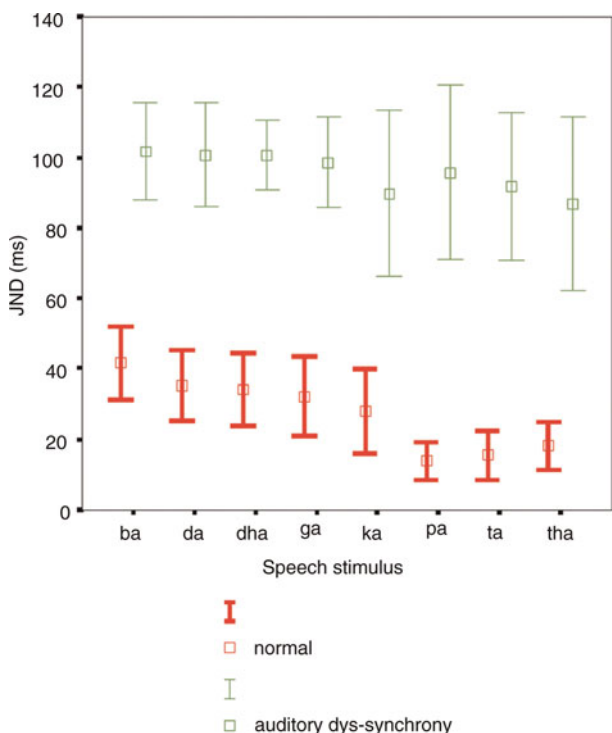


FIG. 4

Just noticeable difference time (JND) for the transition duration of each of the eight consonant–vowel syllables, for the two subject groups. Central squares indicate means; error bars show 95 per cent confidence intervals. AD = auditory dys-synchrony.

in the two groups’ just noticeable difference times, for all sound signals. Figure 4 shows the means and 95 per cent confidence intervals for the observed just noticeable difference times.

Experiment two: individual data

Figure 5 shows individual subjects’ speech identification scores for unmodified and modified (i.e. lengthened transition duration) speech sounds.

Speech identification scores varied between 0 and 87 per cent for the unmodified speech sounds. When presented with the unmodified speech sounds, 11 auditory dys-synchrony subjects had a speech identification score of 0 per cent, while only four auditory dys-synchrony subjects had speech identification scores of more than 50 per cent.

Subjects’ speech identification scores for the modified speech sounds ranged between 0 and 100 per cent. When presented with the modified speech sounds, six auditory dys-synchrony subjects had a speech identification score of 100 per cent.

Experiment two: auditory dys-synchrony group data

Figure 6 shows mean speech identification scores for unmodified and TD lengthened speech sounds. Data from all the subjects have been combined for each speech sound and shown in the figure.

Table II gives a group stimulus-response matrix for unmodified speech sound identification by the auditory dys-synchrony subjects. Each stimulus was presented 10 times, resulting in a total of 300 presentations of each stimulus for the 30 auditory dys-synchrony subjects. However, the row total for each stimulus was less than 300, as on some occasions the subjects could not identify the stimuli and reported hearing only noise. In this stimulus-response matrix, the number in each cell represents the number of times

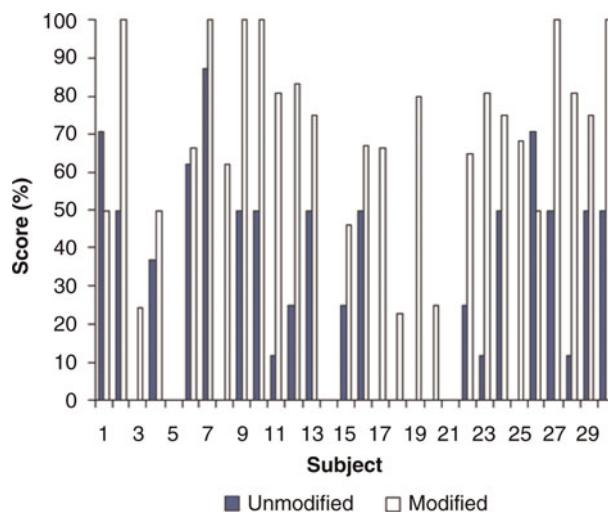


FIG. 5

Speech identification scores of individual subjects for both unmodified and modified (i.e. lengthened transition duration) speech stimuli.

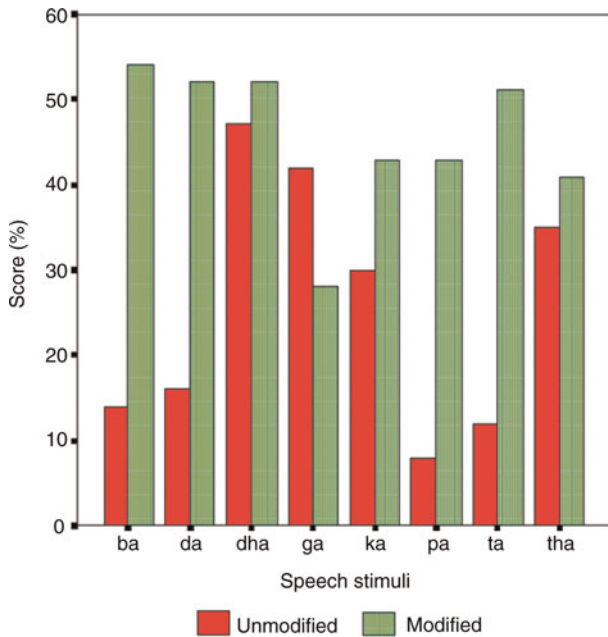


FIG. 6

Mean speech identification scores for unmodified and modified (i.e. lengthened transition duration) speech stimuli, for all auditory dys-synchrony subjects combined.

the speech sound shown in the row heading was identified as the speech sound shown in the column heading. The number of correct responses (for all sounds) can be obtained by summing the numbers occurring along the main diagonal of the table (i.e. from the top left data cell to the bottom right data cell). The matrix reveals that auditory dys-synchrony subjects correctly identified /da/ more than any other sound. The syllable /ga/ was the next best identified. However, identification scores did not exceed 50 per cent for any of the unmodified speech sounds presented. The stimulus-response matrix shows little consistent grouping of results among phoneme categories. The two exceptions were (1) the syllable /ba/, which was frequently confused with its unvoiced cognate /pa/, and (2) the syllable /da/, which was confused with /ga/. The syllables /ta/ and /pa/ were only rarely identified correctly.

Table III gives a group stimulus-response matrix for modified speech sound identification, for the auditory

dys-synchrony subjects. This Table includes only those data for the just noticeable difference modification that resulted in near-normal perception of the consonant–vowel syllable. The data in this stimulus-response matrix should not be directly compared with those in Table II, as the number of presentations of unmodified and modified speech sounds was different. It is noteworthy that confusion between speech sounds was greatly reduced when the transition duration was lengthened. There was greater consistency in the pattern of errors between speech sounds. Lengthening of the transition duration reduced ‘place’ confusion between the phonemes. Confusion of /ba/ with /pa/ was reduced but not eliminated. Also, /dha/ was frequently confused with /da/.

Sequential information transfer analysis²⁸ was performed on the group data for each experimental condition, to assess the amount of information transfer from stimulus to response for each set of phonetic features. Sequential information transfer analysis was performed using the Feature Information Xfer software package (developed by the University College of London linguistics department), which uses the procedure described by Wang and Bilger.²⁸ In sequential information transfer analysis, features go through a number of iterations. In the present study, there were two iterations. The first iteration was the same as the information transmission analysis described by Miller and Nicely,²⁹ wherein information transmitted for each feature is calculated. In the subsequent iterations, the feature with the highest percentage of information transmitted in the previous iteration is held constant and is partialled out. Thus, sequential information transfer analysis helps to estimate the redundancy of a specific feature in its contribution to the perception of monosyllables.

Table IV shows the features assigned to the eight speech sounds used in the study. These eight speech sounds were basically differentiated by their voicing and placement of articulation, as all eight were stop consonants.

Table V shows the relative information transmitted by each feature, expressed as bits, for the unmodified and modified speech sounds. Maximum information

TABLE II
STIMULUS RESPONSE MATRIX FOR IDENTIFICATION OF UNMODIFIED SPEECH SOUNDS BY AUDITORY DYS-SYNCHRONY SUBJECTS

Stimulus	Response							
	ba	da	dha	ga	ka	pa	ta	tha
ba	42 (14)		12 (4)	12 (4)	12 (4)	96 (32)	60 (20)	24 (8)
da		48 (16)	60 (20)	24 (8)	12 (4)	12 (4)	36 (12)	
dha	12 (4)		143 (47)	63 (21)				
ga	12 (4)	60 (20)	24 (8)	128 (43)	15 (5)	12 (4)		
ka		48 (16)	12 (4)	24 (8)	100 (33)	12 (4)	24 (8)	24 (8)
pa				17 (6)	24 (8)	24 (8)	48 (16)	72 (24)
ta	12 (4)	48 (16)	12 (4)	36 (12)		12 (4)	36 (12)	50 (17)
tha		31 (10)			24 (8)	42 (14)	12 (4)	105 (35)

Bold numbers indicate correct responses. Data represent combined responses for all subjects (percentage correct scores, rounded to nearest integer).

TABLE III
STIMULUS RESPONSE MATRIX FOR IDENTIFICATION OF MODIFIED* SPEECH SOUNDS BY AUDITORY DYS-SYNCHRONY SUBJECTS

Sound	Response							No response	Total number of presentation	
	ba	da	dha	ga	ka	pa	ta			tha
ba	144 (55)	12 (4)	12 (4)	5 (2)		36 (14)	12 (4)	11 (4)	32 (12)	264
da		139 (52)	52 (20)	12 (4)				12 (4)	49 (18)	264
dha			95 (52)					3 (1)	82 (45)	180
ga		36 (19)	24 (12)	55 (29)	18 (9)				59 (31)	192
ka		12 (5)			96 (44)	24 (10)	14 (6)		74 (34)	220
pa				12 (5)		115 (44)	24 (8)	24 (8)	89 (34)	264
ta		24 (9)	24 (9)			12 (5)	141 (51)	13 (5)	62 (22)	276
tha					4 (2)		12 (5)	90 (42)	110 (51)	216

Bold numbers indicate correct responses. Data represent combined responses for all subjects (Values in parenthesis indicate percentage correct scores, rounded to nearest integer). *Lengthened transition duration.

that can be transmitted for eight stimuli is three bits (i.e. the maximum information which can be transmitted is $2n$, where n = number of stimuli). In Table V, the numbers in the second and third columns represent the proportion of input information transmission. Zero indicates no transmission of that particular feature, while a value of 1 indicates maximum transmission of information. Numbers in the fourth column represent the total information transmitted, which can range between 0 and 3 bits. In general, the patterns seen in the total transferred information and in the information transfer for each phonetic feature are similar to the pattern found in the stimulus-response matrices shown in Tables II and III and a simple tally of consonants correctly identified.

In individuals with auditory dys-synchrony, lengthening the transition duration appears to result in better transmission of placement information, compared with voicing information.

Discussion

The important findings of this study were: (1) individuals with auditory dys-synchrony have severely impaired temporal processing abilities (as shown by their long just noticeable difference times), and (2) lengthening of speech signal transition duration significantly improved speech perception in such individuals.

The existence of temporal processing deficits in individuals with auditory dys-synchrony has been well documented using non-speech signals.¹⁰⁻¹² Data from the present study indicate that individuals with auditory dys-synchrony have severe difficulties in differentiating

speech segments with altered transition durations, as indicated by their long just noticeable difference times. The just noticeable difference times of auditory dys-synchrony subjects were approximately three to four times longer than those of normal hearing listeners. This may mean that such individuals will have difficulty discriminating between speech sounds that differ only (or mainly) in their temporal characteristics.

Kraus *et al.*¹³ reported a subject with auditory dys-synchrony who had marked difficulty in discriminating between speech sounds of differing spectral onset. Transitions comprise rapid spectral changes occurring at the onset of a speech stimulus. Results from the present study, together with those of Kraus *et al.*,¹³ provide evidence that individuals with auditory dys-synchrony have difficulties in processing temporal and spectral information presented at the onset of speech stimuli.

In the present study, confusion matrices and sequential information transfer analysis showed that lengthening the transition duration of speech stimulus resulted in better transfer of information. However, the results of sequential information transfer analysis should be interpreted with caution. Such analysis is robust when the stimuli occur with equal frequency; however, this could not be ensured in the present study. The number of stimuli was different for different speech sounds, depending on whether they were identified in the unmodified speech condition or not. Obviously, speech syllables which were identified in their unmodified condition were not included in subsequent experiments. This limitation should be kept in mind when

TABLE IV
FEATURES ASSIGNED TO CONSONANT-VOWEL SYLLABLES USED

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

+ = present; - = absent; b = bilabial; a = alveolar; d = dental; v = velar

TABLE V
RELATIVE INFORMATION TRANSMITTED BY EACH SYLLABLE FEATURE

Speech sound	Voicing (bits)	Place (bits)	Total transmitted information (bits)
Unmodified	0.129	0.186	0.812
Modified*	0.464	0.514	1.535

*Lengthened transition duration.

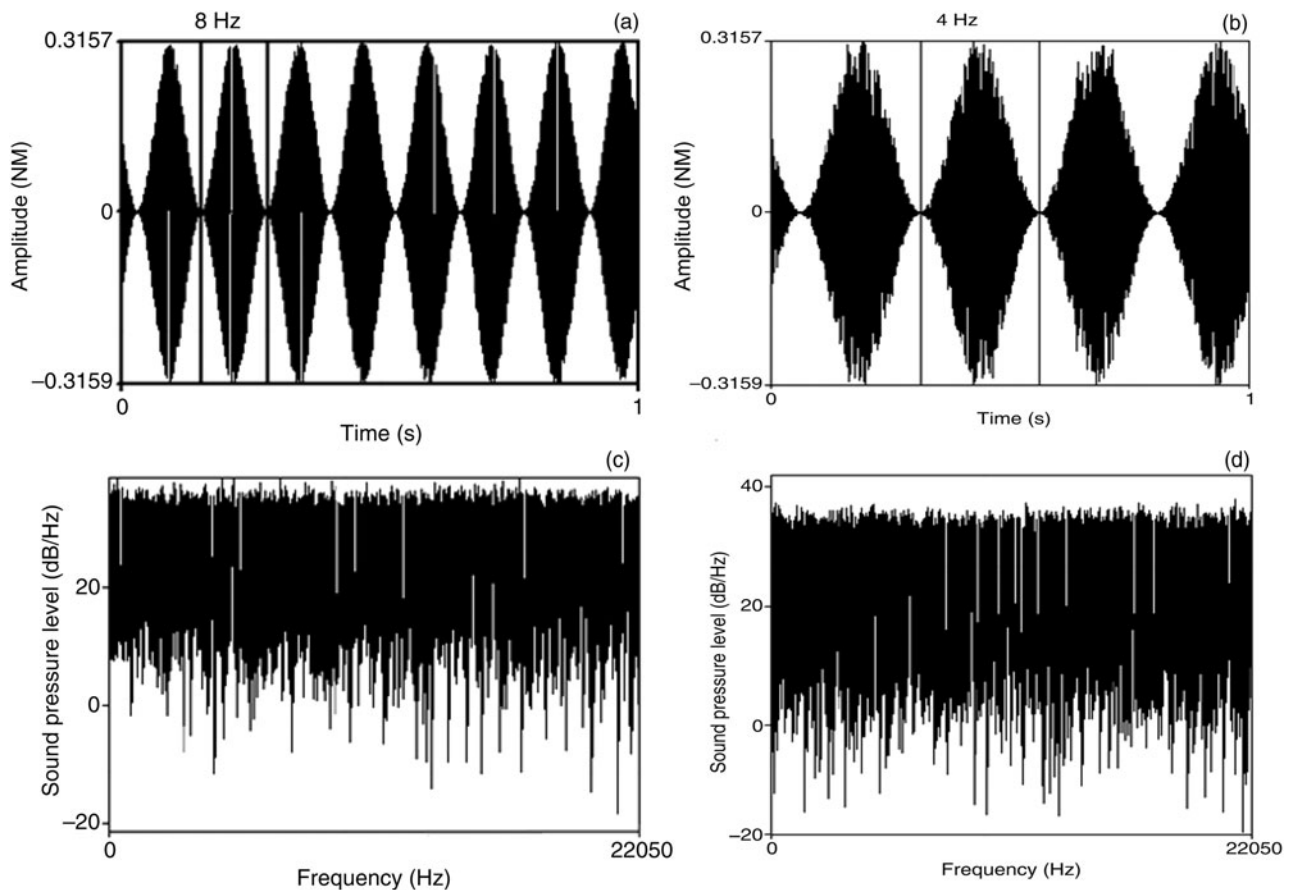


FIG. 7

Frequency spectra and waveforms for unaltered (a and c) and lengthened (b and d) white noise, respectively.

interpreting the results of sequential information transfer analysis.

It has been shown that individuals with auditory dys-synchrony have difficulty in processing short duration sounds.¹¹ Format transitions are short (<50 ms), but they are important acoustic cues for the placement of articulation in stop-vowel syllables. Lengthening of the transition duration of a speech segment may enhance an auditory dys-synchrony individual's perception of this speech segment, but the mechanism is unclear. Similar results have been reported in language learning impaired children.^{30,31} Tallal *et al.*^{30,31} have reported improved speech perception in such children when speech stimuli are modified to increase the duration of format transitions and the modulation depth.

- **Auditory dys-synchrony is characterised by disrupted activity of the auditory nerve in the presence of normal or near-normal functioning of the cochlear outer hair cells**
- **Speech perception is adversely affected**
- **Conventional amplification is of little help**
- **Modifying the temporal structure of speech stimuli may improve speech perception in individuals with this disorder**

We hypothesise that lengthening the transition duration of a speech signal reduces the modulation frequency without altering the modulation depth or the overall spectrogram of the stimulus. This hypothesis was tested by lengthening sinusoidally amplitude modulated white noise by a factor of two. This resulted in a reduction of modulation frequency but no change in modulation depth or spectrogram. Figure 7 shows the spectrogram and waveform of unaltered and lengthened white noise. Data from our laboratory, as well as from others, suggest that individuals with auditory dys-synchrony have difficulty in processing high modulation frequencies.^{11–13} Reduction in modulation frequency (by lengthening the speech signal transition duration) augments such individuals' speech perception, as their modulation detection is better at low modulation frequencies than at high modulation frequencies.

Conclusion

The present study provides indirect evidence that hearing aids which process speech to enhance certain critical short events may be beneficial to persons with auditory dys-synchrony. One such factor is enhancement (lengthening) of transition duration which seems to improve perception of both manner and place cues in individuals with auditory dys-synchrony. Hence, new types of hearing aid which augments/

modifies temporal information of speech may improve speech recognition in patients with auditory dys-synchrony.

Acknowledgement

The authors would like to thank the authorities of the All India Institute of Speech and Hearing, Mysore, for providing the infrastructure for the study. This study forms part of the first author's doctoral thesis.

References

- 1 Starr A, Picton TW, Sininger Y, Hood L, Berlin CI. Auditory neuropathy. *Brain* 1996;**119**:741–53
- 2 Kumar UA, Jayaram M. Prevalence and audiological characteristics in individuals with auditory neuropathy/auditory dys-synchrony. *Int J Audiol* 2006;**45**:360–6
- 3 Berlin C, Hood L, Morlet T, Rose K, Brashears S. Auditory Neuropathy/Dys-synchrony: Diagnosis and Management. *Ment Retard Dev Disabil Res Rev* 2003;**9**:225–31
- 4 Rance G, Beer DE, Cone-Wesson B, Shepherd RK, Dowell RC, King AM *et al*. Clinical findings for a group of infants and young children with auditory neuropathy. *Ear Hear* 1999;**20**:238–52
- 5 Starr A, Michalewski HJ, Zeng FG, Brooks SF, Linthicum F, Kim CS *et al*. Pathology and physiology of auditory neuropathy with a novel mutation in the MPZ gene. *Brain* 2003;**126**:1604–19
- 6 Salvi RJ, Wang J, Ding D, Stecker N, Arnold S. Auditory deprivation of the central auditory system resulting from selective inner hair cell loss: animal model of auditory neuropathy. *Scand Audiol Suppl* 1999;**51**:1–12
- 7 Starr A, Sininger YS, Praat H. Varieties of auditory neuropathy. *J Basic Clin Physiol Pharmacol* 2000;**11**:215–29
- 8 Starr A, Picton TW, Kim R. *Pathophysiology of Auditory Neuropathy*. Canada: Singular Publishing, 2001
- 9 Rance G, McKay C, Grayden D. Perceptual characterization of children with auditory neuropathy. *Ear Hear* 2004;**25**:34–46
- 10 Kumar UA, Jayaram M. Auditory processing in individuals with auditory neuropathy. *Behav Brain Funct* 2005;**1**:8
- 11 Zeng FG, Kong YY, Michalewski HJ, Starr A. Perceptual consequences of disrupted auditory nerve activity. *J Neurophysiol* 2005;**93**:3050–63
- 12 Zeng FG, Oba S, Garde S, Sininger Y, Starr A. Temporal and speech processing deficits in auditory neuropathy. *NeuroReport* 1999;**10**:3439–445
- 13 Kraus N, Bradlow AR, Cheatham MA, Cunningham J, King CD, Koch CD. Consequences of neural asynchrony: a case of auditory neuropathy. *J Assoc Res Otolaryngol* 2000;**1**:33–45
- 14 Berlin CI, Hood LJ, Hurely A, Wen H. *Hair Cells and Hearing Aids*. San Diego: Singular Publishing, 1996
- 15 Plomp R. The negative effect of amplitude compression in multichannel hearing aids in the light of the modulation-transfer function. *J Acoust Soc Am* 1988;**83**:2322–7
- 16 Kraus N. *Auditory Neuropathy: an Historical and Current Perspective*. Canada: Singular Publishing, 2001
- 17 Buss E, Labadie RF, Brown CJ, Gross AJ, Grose JH, Pillsbury HC. Outcome of cochlear implantation in pediatric auditory neuropathy. *Otol Neurotol* 2002;**23**:328–32
- 18 Miyamoto RT, Kirk KI, Renshaw J, Hussain D. Cochlear implantation in auditory neuropathy. *Laryngoscope* 1999;**109**:181–5
- 19 Peterson A, Shallop J, Driscoll C, Breneman A, Babb J, Stoeckel R *et al*. Outcomes of cochlear implantation in children with auditory neuropathy. *J Am Acad Audiol* 2003;**14**:188–201
- 20 Shallop JK, Peterson A, Facer GW, Fabry LB, Driscoll CL. Cochlear implants in five cases of auditory neuropathy: post operative findings and progress. *Laryngoscope* 2001;**111**:555–62
- 21 Trautwein PG, Sininger YS, Nelson R. Cochlear implantation of auditory neuropathy. *J Am Acad Audiol* 1995;**11**:309–15
- 22 Vandana. Speech identification test in Kannada, Unpublished independent project. Mysore: University of Mysore; 1998
- 23 Kumar UA, Vanaja CS. Functioning of olivocochlear bundle and speech perception in noise. *Ear Hear* 2004;**25**:142–6
- 24 Jayaram M. Distribution of stuttering in sentences: relationship to sentence length and clause position. *J Speech Hear Res* 1984;**27**:38–42
- 25 Boersm & Weenink. Praat: Doing Phonetics by Computer (version 4.2.31) 2005, retrieved 4 August 2005 from <http://www.praat.org/>
- 26 Moulines E, Lorché J. Non parametric techniques for pitch scale and time scale modification of speech. *Speech Commun* 1995;**16**:175–205
- 27 Tyler RS, Summerfield Q, Wood EJ, Fernandes MA. Psychoacoustic and phonetic temporal processing in normal and hearing impaired listeners. *J Acoust Soc Am* 1982;**72**:740–52
- 28 Wang MD, Bilger RC. Consonant confusions in noise: a study of perceptual features. *J Acoust Soc Am* 1973;**54**:1248–66
- 29 Miller AJ, Nicely PE. An analysis of perceptual confusions among some English consonants. *Journal of Acoustical Society of America* 1955;**27**:338–52
- 30 Tallal P, Miller ST, Bedi G, Byma G, Wang X, Nagarajan S *et al*. Language comprehension in language learning impaired children improved with acoustically modified speech. *Science* 1996;**271**:81–4
- 31 Tallal P, Piercy M. Developmental aphasia: the perception of brief vowels and extended stop consonants. *Neuropsychologia* 1975;**13**:69–74

Address for correspondence:

Dr U Ajith Kumar,
Associate Professor,
Department of Audiology and Speech Language Pathology,
Kasturba Medical College (a unit of Manipal University),
Mangalore, India

E-mail: ajithkumar18@gmail.com

Dr A Kumar takes responsibility for the integrity of the content of the paper
Competing interests: None declared
