Proc. National Conference on Biomedical Engg., Roorkee. India, April 21-22, 2000. pp 346-353.

DICHOTIC PRESENTATION WITH INTER-AURAL SWITCHING FOR REDUCING THE EFFECT OF TEMPORAL MASKING DUE TO SENSORINEURAL HEARING LOSS

D. S. Jangamashetti* and P. C. Pandey+
*Research Scholar, +Associate Professor
Department of Electrical Engineering
Indian Institute of Technology, Bombay

Synopsis

Sensorineural hearing loss, is associated with increased forward and backward masking of weak acoustic sub-segments by strong ones, thereby degrading the reception of important cues for identification of consonants. Employment of inter-aural switching, by splitting the speech temporally and presenting alternate segments to the two ears may help in reducing the effect of temporal masking. A symmetrical inter-aural switching period of 20 ms was used for temporal splitting of the speech signals. Effect of different types of fading functions and overlap (duty cycle greater than 50%) in reducing the perception of temporal gaps due to switching and in improving speech intelligibility was studied. Test stimuli consisted of twelve English nonsense syllables in vowel-consonant-vowel context with vowel /a/. Four duty cycles of 100%, 75%, 70%, and 60% were used along with step transition for switching. Listening tests were conducted on three normal hearing subjects, with sensorineural loss simulated by adding broadband noise to the speech signal at different signal-to-noise ratios (∞, 6, 3, 0, and -3 dB). Test results are analyzed by comparing the recognition scores for unprocessed and processed speech, and also the relative transmission of information for consonantal features. It has been observed that, the inter-aural switching with 70% duty cycle, for dichotic presentation helps in improving the speech perception.

A. Introduction

Sensorineural loss is characterized by elevated thresholds, loudness recruitment, reduced frequency and temporal resolution, and increased spectral and temporal masking. Various compression techniques have been employed to improve the speech perception, degraded due to loss of dynamic range due to loudness recruitment [1,2]. Masking is a phenomenon where presence of one signal component results in in-audibility of the neighboring signal component. The masking takes place primarily, at the level of peripheral auditory system. In speech perception, the information received from both the ears gets integrated. Hence splitting of information in speech signal for presenting signals to the two ears, in some sort of a complimentary fashion, may help in reducing the effect of masking. Presentation of different signals to the two ears is known as dichotic binaural presentation. For reducing the effect of increased spectral masking, spectral splitting of speech signal for dichotic binaural presentation has been investigated and real time processing and implementations have been carried out [3,4,5]. Increased temporal masking results in increased forward and backward masking of weak acoustic segments by strong ones. This degrades the reception of important cues for identification of consonants, and results in degradation of speech perception. For reducing the effect of increased temporal masking, consonantal enhancement based on the properties of "clear speech" has been investigated [6,7,8]. However, no implementation of these schemes has been reported.

Splitting the speech temporally into a number of segments and presenting the alternate segments to the two ears may help in reducing the effect of increased temporal masking [3]. Lunner et al. [3] investigated splitting of speech signal into odd and even bands (with 700 Hz bandwidth) for binaural dichotic presentation, with a symmetrical inter-aural switching of odd and even bands with a period of 20 ms. They reported an improvement due to spectral splitting. However, inter-aural switching of the bands did not result in any improvement in speech perception. The sound quality was found to deteriorate, possibly due to perception of a small gap at switching transitions.

Provision of a certain overlap during inter-switching (symmetrical inter-aural switching with duty cycle >50%) may help in reducing the perception of temporal gaps due to switching, and may thereby increase the speech intelligibility for persons with sensorineural hearing loss and for normal hearing persons under adverse listening conditions. The objective of the research reported here was to split the speech using symmetrical inter-aural switching along with overlap for binaural dichotic presentation and to study the effect of various duty cycles and "fading functions" in increasing the speech intelligibility. Fading functions with step and trapezoidal transitions were investigated and experimentally evaluated by conducting listening tests on normal hearing subjects with simulated sensorineural hearing loss. Processing of the signal, experimental evaluation, and results are discussed in the following sections.

B. Processing

Figures 1(a) and 1(b) represent temporal processing with step transitions. The outputs $s_1(n)$ and $s_2(n)$ of temporal processing were obtained by multiplying input $s(n)$ with fading functions $w_1(n)$ and $w_2(n)$ respectively. The two windows have a symmetrical overlap, each having a duty cycle $d = L/N$, where N = inter-aural switching duration and L = "on" period.

$$w_1(n) = \begin{cases} 1, & 0 \le (n)_N \le L-1 \\ 0, & \text{otherwise} \end{cases}$$

$$w_2(n) = \begin{cases} 0, & L - N/2 + 1 \le (n)_N \le N/2 \\ 1, & \text{otherwise} \end{cases}$$

For an inter-aural switching period of N samples with 50% duty cycle, alternate segments of $N/2$ samples are presented to the ears such that adjacent segments get presented to the different ears, i.e. while first $N/2$ samples of signal are presented to left ear, right ear is relaxed and during next $N/2$ samples of signal presentation to the right ear, left ear is relaxed. This scheme will help in reducing the temporal masking. However, it is likely to introduce

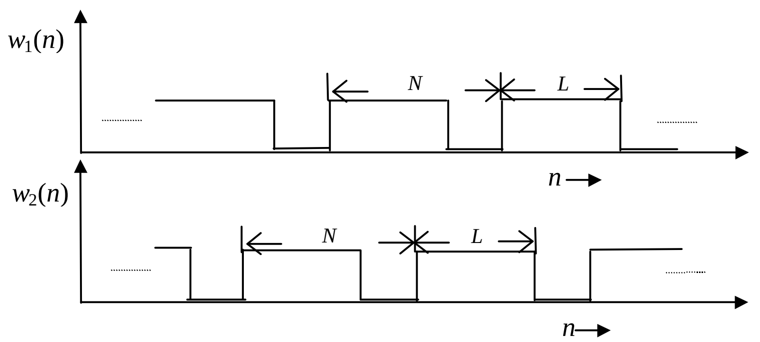

FIG. 1(a). Overlapping inter-aural switching with step

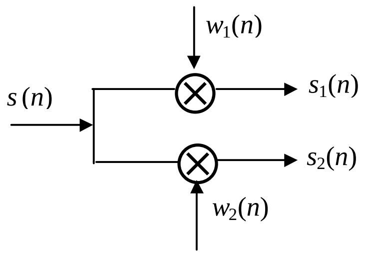

FIG. 1(b). Temporal splitting of the signal.

perceptual gaps around the transitions, related to inter-aural switching. Splitting the signal with duty cycles greater than 50% will provide certain overlap between the two presentations and will help in reducing the perception of inter-aural switching. Instead of using sharp transitions at switching, smooth variation may help in improving sound quality by avoiding the occurrence of high frequency components due to abrupt transition. Hence temporal processing scheme includes: (i) splitting the signal with different duty cycles with step transition, i.e. using fading function shown in Fig. 1(a) and (ii) splitting the signal with different duty cycles with trapezoidal transition. Duty cycles greater than 50% and trapezoidal transitions will reduce the perception of inter-aural switching, but will leave less time available for relaxation of the auditory sensors and nerve fibers. In the present work, implementation and testing of temporal splitting with step transition has been carried out and the results are analyzed for duty cycles of 100, 75, 70, and 60%.

C. Methods and Materials

Three normal hearing subjects (two males and one female) participated in the study. The subjects had hearing thresholds less than 20 dB HL in the frequency range 125 Hz to 6 kHz and their ages were in the range 20–31 years. Twelve English consonants /p, b, t, d, k, g, m, n, s, z, f, v/ were used in the vowel-consonant-vowel (VCV) context with vowel /a/. Nonsense syllables were used to minimize the contribution of linguistic factors. These stimuli were acquired at 10 k samples/s, using a microphone, a preamplifier, a 7th order lowpass filter with a cut-off frequency of 4.8 kHz, and a 16-bit A/D converter of TI / TMS320C25 DSP board, interfaced to a PC.

Speech discrimination performance deteriorates at low (<35 dB) and high (>90 dB) presentation levels [9,10,11]. In our experiment, stimuli were presented at most comfortable listening level for the individual subjects and were in the range 70–75 dB SPL. Simulation of sensorineural hearing loss in normal hearing persons can be done by adding the broadband noise to the speech stimuli [12,13]. In our study, signal-to-noise ratios (SNR) of ∞, 6, 3, 0, -3 dB were used to simulate sensorineural hearing loss of varying degrees. The noise used is broad band Gaussian noise from function generator HP 33120A.

Figure 2 shows the experimental set-up used for conducting the listening tests. The stimuli were presented at 10 k samples/s through the two D/A ports of PCL-208 data acquisition card. The signals after passing through a 7th order smoothing low pass filter with cutoff frequency 4.8 kHz and an audio amplifier, were given to the two head phones (Telephonics TDH-39P).

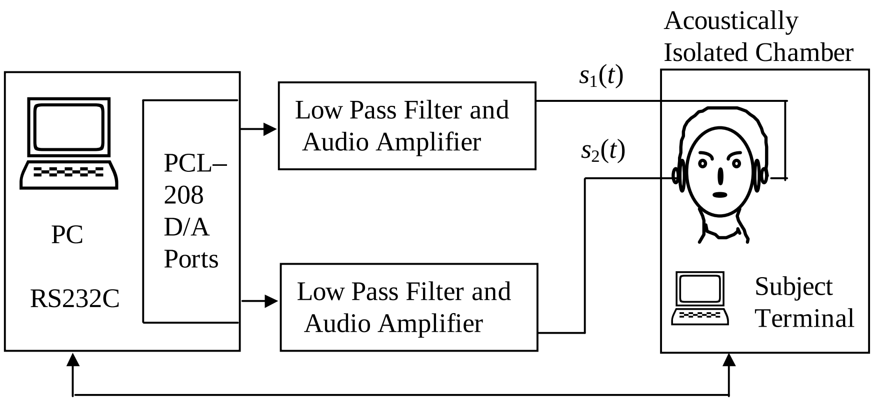

FIG. 2. Listening test set–up.

The computerized test administration system consisted of a PC interfaced through RS232C serial port, to the subject terminal (VT-220) which is placed in an acoustically isolated chamber. PC was used for controlling the entire test and the subject terminal was used for displaying the response choices and obtaining the subjects responses from its keyboard.

A program written in C for temporal splitting provides facility to select desired inter-aural switching period, duty cycle, and transition duration at switching (for fading function with trapezoidal transition). In the sounds processed by temporal splitting, a certain amount of inter-aural switching is perceived and it results in loss of speech quality. The perception of the switching reduces with increase in duty cycle. As compared to step transition, trapezoidal transition generally results in lesser loss of quality. During transition period, either linear variation of the sample value or logarithmic variation of the samples can be used. Stimuli processed with linear sample variation were better heard than those obtained with logarithmic variation. Step transition provides larger interval for relaxation, and therefore may be more effective in reducing the effect of temporal masking, although it may result in degradation in the reception of some of the spectral characteristics.

To evaluate the inter-aural switching scheme with step transition, and to find the near-optimal duty cycle, listening tests were conducted. Processed and unprocessed stimuli were added with broadband noise at different SNR conditions to simulate sensorineural hearing loss in normal hearing subjects. There were 20 (4 duty cycles x 5 SNR) test conditions.

D. Results and Discussion

Test results are obtained in the form of confusion matrix with rows and columns representing stimuli and responses respectively. From the confusion matrices, percentage recognition scores were calculated. Table 1 shows average percentage recognition scores of five stabilized sets for unprocessed signal and % relative improvement for processed signal with different duty cycles for each subject. It can be observed that, for unprocessed stimuli (100% duty cycle), scores are perfect under no masking noise and decrease as the level of noise increases. For processed stimuli, under no masking condition the scores decrease as the duty cycle decreases. This indicates that, temporal splitting degrades speech perception under normal listening conditions. However, when noise is present, the scores for the processed stimuli are higher. All the three subjects have shown improvement in score with processing and maximum scores are seen for 70% duty cycle. For higher levels of noise, there is more improvement in recognition score.

Information transmission analysis [14] was carried out to obtain the relative information transmitted for features of voicing, place, manner, nasality, frication, and duration. Table 2 shows the percentage overall information transmitted for 100% duty cycle and relative improvement for 75, 70, and 60% duty cycles, for all subjects. Table 3 gives information transmitted and % relative improvement, respectively for unprocessed and processed speech, for voicing, place, and duration features for subject S1. For voicing, relative information transmission is perfect for ∞, 6, and 3 dB masking conditions. For 0 dB and -3 dB conditions, relative information transmission decreases for unprocessed stimuli whereas it increases for processed stimuli. There is improvement for place feature, and duration feature at all SNR conditions for processed stimuli. For place feature, a maximum relative improvement of up to 54% is noted at -3 dB SNR condition whereas, for duration feature maximum relative improvement of up to 45% is observed at 0 dB SNR condition. It is to be noted that with the decrease in SNR, the perception of the duration feature gets degraded most. Improvement in its reception by inter-aural switching indicates that processing has reduced the effect of temporal masking.

E. Conclusion

From the results it can be seen that there is an improvement in recognition scores and relative information transmission, and improvements depend on duty cycle of the inter-aural switching function. Duty cycle of 70% gives maximum improvement. Use of fading function with trapezoidal variation may help in further improvement by reducing occurrence of undesired frequency components due to step transitions. Further, combination of temporal splitting with spectral splitting should be investigated.

TABLE 1. Recognition scores (%) for unprocessed speech (Su) and relative improvement (%) for processed speech (Sp) under five SNR conditions for three subjects.

Sub-ject	∞ SNR				6 dB SNR				3 dB SNR				0 dB SNR				-3 dB SNR			
	Su	Sp			Su	Sp			Su	Sp			Su	Sp			Su	Sp		
		%Duty cycle				%Duty cycle				%Duty cycle				%Duty cycle				%Duty cycle		
		75	70	60		75	70	60		75	70	60		75	70	60		75	70	60
S1	100	-2	-3	-6	91	2.1	5.5	4.4	95	-1	3.1	2.1	91	3.3	3.3	1.1	80	6.2	10	7.5
S2	98	-2	-2	-2	81	3.7	3.7	6.1	78	1.3	6.4	3.8	71	4.2	7.0	8.4	67	-1.5	12	4.5
S3	100	-2	-2	-3	95	4.2	4.2	2.1	95	2.1	3.1	3.1	93	-1	-1	2.1	86	3.4	1.2	5.8
Avg.	99	-2	-2	-3	89	3.4	4.5	3.4	89	0	3.4	2.2	84	2.4	3.6	2.4	78	1.3	6.4	5.1

TABLE 2. Overall information transmitted (%) for unprocessed speech (Su) and relative improvement (%) for processed speech (Sp).

Sub-ject	∞ SNR				6 dB SNR				3 dB SNR				0 dB SNR				-3 dB SNR			
	Su	Sp			Su	Sp			Su	Sp			Su	Sp			Su	Sp		
		%Duty cycle				%Duty cycle				%Duty cycle				%Duty cycle				%Duty cycle		
		75	70	60		75	70	60		75	70	60		75	70	60		75	70	60
S1	100	-2	-4	-5	90	2.2	5.5	4.4	93	1.1	4.3	2.2	90	2.2	2.2	0	82	3.6	6.1	6.1
S2	94	2	1	2	86	0	-2.3	-1.1	81	2.5	3.7	1.2	80	0	2.5	-1.2	77	-5	5.2	-2.6
S3	99	-1	-1	-3	96	3.1	3.1	0	95	2.1	3.2	2.1	92	0	0	0	90	0	-1.1	1.1
Avg.	97	0	-1	-1	91	1.1	2.2	1.1	90	1.1	3.3	1.1	87	1.1	2.3	0	83	0	4.8	1.2

TABLE 3. Relative Information transmitted (%) for unprocessed speech (Su) and relative improvement (%) for processed speech (Sp) for features of voicing, place and duration for subject S1. (F–Feature, Vo–Voicing., Pl–Place, Du–Duration)

F	∞ SNR				6 dB SNR				3 dB SNR				0 dB SNR				-3 dB SNR			
	Su	Sp			Su	Sp			Su	Sp			Su	Sp			Su	Sp		
		%Duty cycle				%Duty cycle				%Duty cycle				%Duty cycle				%Duty cycle		
		75	70	60		75	70	60		75	70	60		75	70	60		75	70	60
Vo	100	0	0	0	100	0	0	0	100	0	0	0	97	0	3.1	3.1	94	3.2	3.2	6.4
Pl	100	-8	-12	-17	76	0	12	12	81	0	12.3	3.7	70	16	11.4	0	41	29	54	46
Du	100	-19	-19	46	76	0	12	4	83	-1.2	16	-4.8	60	31.7	45	5	43	23	28	7

References

1. J. N. van Dijkhuizen, J. M. Festen, and R. Plomp, "The effect of frequency-selective attenuation on the speech-reception threshold of sentences in conditions of low-frequency noise," *J. Acoust. Soc. Amer.*, vol. 56, no. 5, pp.1601-1611, 1991.
2. J. C. Tejero, S. Bernal, J. A. Hidaldo, J. Fernandez, R. Urquiza, and A. Gago, "A digital hearing aid that compensates loudness for sensorineural hearing impairments," *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing* (*ICASSP*), Málaga, 1995, pp. 2991-2994.
3. T. Lunner, S. Arlinger, and J. Hellgren, "8-channel digital filter bank for hearing aid use: preliminary results in monaural, diotic and dichotic modes," *Scand. Audiol.* Suppl. Vol. 38, pp. 75-81, 1993.
4. D. S. Chaudhari and P. C. Pandey, "Dichotic presentation of signal using critical filter bank for bilateral sensorineural hearing impairment," *Proc. of 16th Int. Congress on Acoust.* (*ICA*), Seattle, Wash., U.S.A., 1998, pp. 213-214.
5. D. S. Chaudhari and P. C. Pandey, "Dichotic presentation of speech signal with critical band filtering for improving speech perception," *Proc. ICASSP'98,* Seattle, Wash., U.S.A., 1998, Paper AE 3.1.
6. M. A. Picheny, N. I. Durlach, and L. D. Braida, "Speaking clearly for the hard of hearing II: Acoustic characteristic of clear and conversational speech," *J. Speech Hear. Res.*, vol. 29, pp. 434–446, 1986.
7. T. G. Thomas, P.C. Pandey, and S. D. Agashe, "Effect of consonantal intensity and duration modification on speech perception by listeners with simulated hearing impairment," *J. Acoust. Soc. India*, XXIV, VI-4.1-4.5, 1995.
8. T. G. Thomas, P. C. Pandey, and S. D. Agashe, "On the importance of consonant-vowel intensity ratio in speech enhancement for the hearing impaired," *Proc. Int. Conf. Biomed. Engg,* Hong Kong, 1994, pp. 181-184.
9. S. Gordon-Salant, "Effects of acoustic modification on consonant recognition by elderly hearing-impaired subjects," *J. Acoust. Soc. Amer.*, vol. 81, pp. 1199–1202, 1987.
10. C. Simon, "On the use of comfortable listening levels in speech experiments," *J. Acoust. Soc. Amer.*, vol. 64, no. 3, pp. 744–750, 1978.
11. M. F. Dorman and K. Dougherty, "Shifts in phonetic identification with changes in signal presentation level," *J. Acoust. Soc. Amer*, vol. 69, no. 5, pp. 1439–1440, 1981.
12. L. E. Humes, D. D. Dirks, T. S. Bell, and G. E. Kincaid, "Recognition of nonsense syllables by hearing-impaired listeners and by noise-masked normal hearers," *J. Acoust. Soc. Amer.*, vol. 81, no. 3, pp. 765–773, 1987.
13. J. R. Dubno and A. B. Schaefer, "Comparison of frequency selectivity and consonant recognition among hearing-impaired and masked normal-hearing listeners," *J. Acoust. Soc. Amer.*, vol. 91, no. 4, Pt.1, pp. 2110–2121, 1992.
14. G. A. Miller and P. E. Nicely, "Analysis of perceptual confusions among some English consonants," *J. Acoust. Soc. Amer.*, vol. 27, no. 2, pp. 338-352, 1955.