Int. Symposium on Frontiers of Research on Speech and Music 2008 (FRSM 2008), Kolkata, India, pp. 101-105.

STUDY OF PERCEPTUAL BALANCE IN COMB FILTER BASED SPECTRAL SPLITTING OF SPEECH SIGNAL TO REDUCE THE EFFECT OF FREQUENCY MASKING

P. N. Kulkarni[1], P. C. Pandey[2], and D. S. Jangamashetti[3]

[1,2]SPI Lab, Department of Electrical Engineering
Indian Institute of Technology Bombay
Powai, Mumbai 400 076, India
E-mail:{pnkulkarni, pcpandey}@ ee.iitb.ac.in
[3]Department of Electrical and Electronics Engineering
Basaveshwar Engineering College
Bagalkot 587102, Karnataka, India
E-mail: dsjee@becbgk.edu

Abstract- Earlier investigations have shown that spectral splitting of speech signal helps in reducing the effect of spectral masking for persons with moderate bilateral sensorineural hearing loss. In spectral splitting, the speech signal is filtered using a pair of linear phase FIR comb filters with complementary magnitude responses for binaural dichotic presentation. The filter responses should be such that perceived loudness remains balanced, especially for spectral components in transition bands which get presented to both the ears. Relationship between scaling factors for a tone presented to the two ears, so that perceived loudness is that of a monaural presentation, is investigated for design of comb filters with improved perceptually balanced response. Results from the listening tests show that sum of the two scaling factors should be constant, indicating that the magnitude response of the comb filters should be complementary on a linear scale.

Keywords- Sensorineural hearing loss, frequency masking, spectral splitting, comb filter, binaural perceptual balance.

1. Introduction

Sensorineural hearing loss occurs when the functioning of the cochlea is affected or when there is dysfunction of the auditory nerve or higher centers in the auditory pathway [1]. It is characterized by elevated hearing thresholds, loudness recruitment and reduced dynamic range, poor frequency and temporal resolution, increased temporal and spectral masking, and degraded speech perception [1]-[4]. Earlier investigations [5]-[11] have shown that splitting the speech signal into different bands and presenting the alternate bands to the left and right ear improved speech perception for persons with moderate bilateral sensorineural hearing impairment, with residual hearing in both the ears. In spectral splitting technique, based on comb filters, a specific spectral component is generally presented to one ear. However, spectral components in the transition bands get presented to both the ears. Hence filter magnitude responses should have sharp transitions. Further, the two filters should have magnitude responses such that a perceptual loudness balance is maintained for spectral components in the transition bands. This necessitates a detailed study of the transition band response of the comb filters, so that perceptual balance is maintained when the same spectral components are simultaneously presented to the two ears. The objective of the present investigation is to study the perceptual balance in binaural listening, i.e., finding different combinations of intensities in the left and right ear which will evoke the same loudness as a monaural presentation.

Experiments were conducted by Whilby *et al*. [12], Hawkins *et al*. [13], Hall and Harvey [14], Scharf [15], [16], Epstein and Florentine [17], to investigate binaural level difference for equal loudness (BLDEL) in normal-hearing and hearing-impaired listeners. Whilby *et al*. investigated BLDEL for normal and hearing impaired listeners using 1 kHz pure tone of 5 and 200 ms duration. The objective of their study was to find BLDEL, using loudness matching procedure: (1) monaural level was fixed and binaural level was varied for equal loudness and (2) binaural level was fixed and monaural level was varied. The fixed level ranged from 10 to 90 dB SL. Their study showed that BLDEL for normal-hearing subjects ranged from 2 to 15 dB, and for hearing impaired listeners it was 1.5 to 12 dB. They found that on an average, BLDEL was independent of duration but depends upon the presentation level. Hawkins *et al.* experimented for BLDEL in normal and hearing impaired listeners, with a presentation level ranging within listener's most comfortable level to discomfort level, using 4 kHz pure tone. They reported that BLDEL for impaired listeners was not significantly different from that for listeners with normal hearing. However, in the investigation by Hall and Harvey, impaired listeners had less BLDEL than normal listeners when the monaural tone was at 70 and 80 dB SPL, while BLDEL was comparable at 90 dB SPL. Studies by Scharf [15], [16] indicate that BLDEL is a function of presentation level and is small at low presentation levels, high at moderate presentation levels, and moderate at high presentation levels.

Marks [18] investigated binaural summation of loudness using pure tone stimuli of frequency 0.1, 0.4, and 1 kHz, on normal hearing subjects. A set of nine sound pressure levels, ranging from $-\infty$ to 50 dB SPL, were used for both left and right ears (a total of 81 combinations of binaural stimuli). Presentation for each of the stimuli was randomized with each stimulus presented twice. Fourteen normal hearing subjects participated in the listening test. Subjects were asked to assign a number appropriate to represent the loudness of the first stimulus and to assign numbers in proportion to the subsequent stimuli. If no sound was heard then subjects assigned zero. Thus number assigned by the subjects gave a quantitative measure of perceived loudness and was referred to as magnitude estimate of loudness. For each test tone frequency, mean magnitude estimate of the loudness as a function of SPL presented to the right ear was plotted, for each of the nine SPLs presented to the left ear. A total of nine contours thus obtained were of similar shape but displaced vertically along loudness axis. This characteristic indicated the linear additivity of the numerical responses.

Cheeran [6] conducted an experiment to find the difference in intensities of monaural and binaural presentations, such that they evoke same perceived loudness. The investigation showed that perceived loudness matches when binaural level was 4-12 dB lower than the monaural level. For 1 kHz tone, the average monaural-binaural level difference was 6.4 dB with a small inter-subject variation. Based on these results, the comb filters were designed with inter-band crossover gain adjusted between 4 to 6 dB lower than the pass band gain [9], [10].

The aim of present investigation is to study the perceptual balance, in binaural listening, as applied to spectral splitting, and hence establishing the relation between the amplitude scaling factors for the left and right ears so that perceptual balance is observed for test tones of different frequencies.

2. Methodology

As shown in Fig. 1, the input signal is scaled by amplitude scaling factor α for presentation to the left ear and by β for presentation to the right ear. Pure tones of frequency 250 Hz, 500 Hz, 1 kHz, and 2 kHz of 1 s duration, sampled at 10 k samples/s are used as the test material. The values of α and β considered for investigation ranges from 0 to 1 with an increment of 0.1.

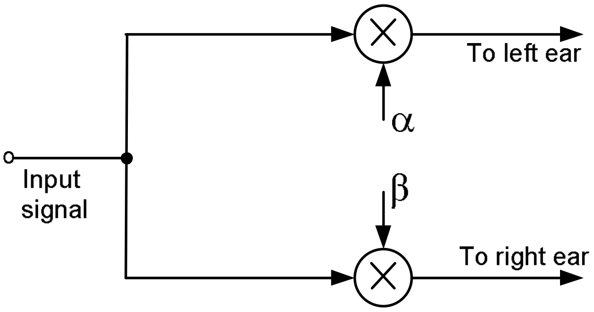

Fig. 1 Scheme for perceptual balance test.

With an objective of finding a relation between α and β for perceptual balance in binaural presentation, listening tests were conducted using three interval, three alternative forced choice (3I, 3AFC) paradigm [19], [20]. Each presentation had three intervals (i.e. reference, test, and reference). Monaural reference sound is presented to left ear. After a 0.5 s silence, the binaural test sound is presented. After a 0.5 s silence, the monaural reference is presented again. For each presentation (i.e. with a binaural sound corresponding to a specific combination α and β), subjects were asked to compare the perceived loudness of binaural sound to that of monaural sound, which is presented before and after the binaural sound, with 0.5 s of silence in between. Subjects were allowed to listen to the sounds as many times as they wished until they finalized their response. The subjects were instructed to respond to each presentation by marking it, on the response sheet, as either L, E, or H depending on whether the perceived loudness of the binaural sound is lower than, equal to, or higher than that of monaural sound.

Four subjects with normal hearing participated in the listening tests. With 15 values of α, 15 values of β, and 4 frequencies, there were 900 presentations for each subject.

3. Results and discussion

For all values of α, range of values of β for which the perceptual balance was observed (i.e. when the response is E) is found. Then for each α, a value of β that evoked perceptual balance is obtained as the mean of this range. Fig. 2 shows the averaged (across the subjects) β as a function of α, for four test frequencies.

Assuming that the loudness is related to a power of sound amplitude (for amplitudes in the range 0-1), and the loudness of binaural sound is a power law summation of the two individual sounds, the scaling factors α and β should have following relationship for perceptual balance,

$$(\alpha)^p + (\beta)^p = 1 \qquad (1)$$

where, p is the power relating amplitude to loudness. Hence,

$$\beta = \left[\left(1-\alpha^p\right)\right]^{1/p} \qquad (2)$$

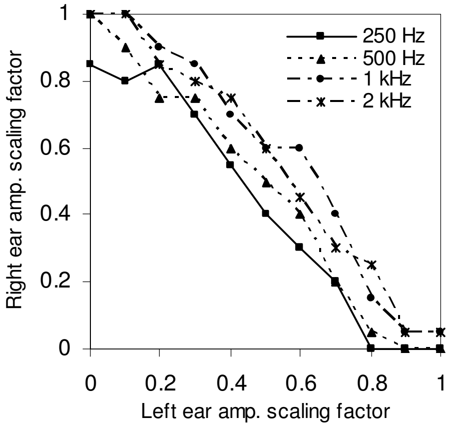

Fig. 2 Averaged (across four subjects) relationship between the left and right ear amplitude scaling factors shown for different frequencies.

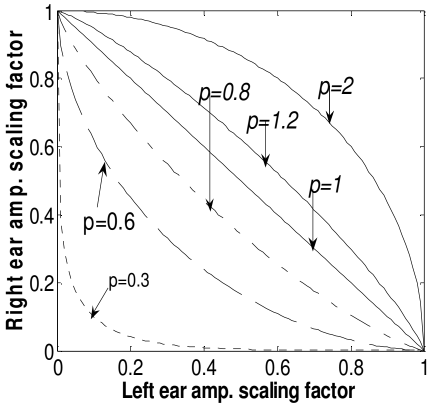

Fig. 3 Power law relation between the left and right ear scaling factors based on (2).

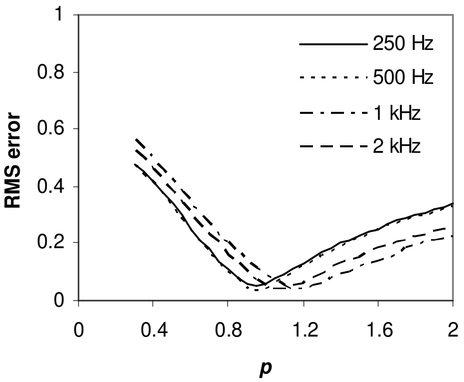

Fig. 4 RMS error in approximation of β as computed from (2) shown as a function of p.

With an aim of finding an approximate fit of the observed values of β, β values were computed for different values of p (0.3, 0.6, 0.8, 1.0, 1.2, and 2) from (2), and these are plotted in Fig. 3. It can be seen that α-β plots for all test tone frequencies in Fig. 2 nearly match the plot for $p = 1$ in Fig. 3. Fig. 4 shows the error in approximation of β, as computed using (2), as a function of p for different test tone frequencies. It can be seen that the error is minimum for $p \approx 1$, indicating approximately linear relation between the two scaling factors. The results obtained are in line with the results reported in our earlier work on perceptual balance [11], in which similar experiments were carried out by selecting α and β on dB scale but using only 1 kHz pure tone.

4. Summary and conclusion

The objective of the present study was to investigate the relationship between the scaling factors for perceptual balance as applied to binaural dichotic splitting of speech to reduce the effect of spectral masking in moderate bilateral sensorineural hearing loss. Listening tests were carried out using 250 Hz, 500 Hz, 1 kHz, and 2 kHz pure tones on normal-hearing subjects. The results show a linear relation between the amplitude scaling factors for left and right ears for perceptual balance, indicating that the magnitude response of the comb filters should be complementary on a linear scale.

References

[1] B. C. J. Moore, *An Introduction to Psychology of Hearing*. 4th ed. London: Academic, 1997.

[2] D. O'Shaughnessy, *Speech Communication: Human and Machine*. New York: Addison-Wesley, 1987.

[3] B. R. Glasberg and B. C. J. Moore, “Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments,” *J. Acoust. Soc. Am.*, vol. 79, pp. 1020-1033, 1986.

[4] A. E. Carney and D. A. Nelson, “An analysis of psychophysical tuning curves in normal hearing and pathological ears,” *J. Acoust. Soc. Am.*, vol. 73, pp. 268-278, 1983.

[5] P. E. Lyregaard, “Frequency selectivity and speech intelligibility in noise,” *Scand. Audiol. Suppl.*, vol. 15, pp. 113-122, 1982.

[6] A. N. Cheeran, "Speech processing with dichotic presentation for binaural hearing aids for moderate bilateral sensorineural loss,” *Ph.D. dissertation*, School of Biosciences and Bioengineering, IIT Bombay. 2005.

[7] D. S. Chaudhari and P. C. Pandey, “Dichotic presentation of speech signal with critical band filtering for improving speech perception,” in *Proc. IEEE / ICASSP 1998*, Seattle, Washington, vol.6, pp. 3601-3604, 1998.

[8] D. S. Jangamashetti, "Binaural dichotic presentation to reduce the effects of temporal and spectral masking due to sensorineural hearing loss,” *Ph.D. dissertation*, Dept. of Elect. Engg, IIT Bombay. 2003.

[9] A. N. Cheeran and P. C. Pandey, “Evaluation of speech processing schemes using binaural dichotic presentation to reduce the effect of masking in hearing-impaired listeners,” in *Proc. 18th International Congress on Acoustics, ICA 2004,* Kyoto, Japan, pp. II-1523 - 1526, 2004.

[10] A. N. Cheeran and P. C. Pandey, “Dichotic presentation for binaural hearing aids using perceptually balanced critical bandwidth based comb filters,” in *Proc. SCI 2003,* Orlando, Florida, pp. IV-360-365, 2003.

[11] P. N. Kulkarni, P. C. Pandey, and D. S. Jangamashetti, “Perceptually balanced filter response for binaural dichotic presentation to reduce the effect of spectral masking,” *J. Acoust. Soc. Am.*, vol. 120, no.5, pp. 3253, 2006.

[12] S. Whilby, M. Florentine, E. Wagner, and J. Marozeau, “Monaural and binaural loudness of 5 and 200 ms tones in normal and impaired hearing,” *J. Acoust. Soc. Am.*, vol. 119, no. 6, pp. 3931-3939, June 2006.

[13] D. B. Hawkins, R. A. Prosek, B. E. Walden, and A. A. Montgomery, “Binaural loudness summation in hearing impaired,” *J. Speech Hear. Res.*, Vol. 30, pp.37- 43, March 1987.

[14] J. W. Hall and A. D. Harvey, “Diotic loudness summation in normal and impaired hearing,” *J. Speech Hear. Res.*, vol. 30, pp. 37- 43, 1985.

[15] B. Scharf, "Binaural loudness summation as a function of bandwidth," in *Proc. The 6th International Congress on Acoustics*, Tokyo, Japan, 1968.

[16] B. Scharf, "Dichotic summation of loudness," *J. Acoust. Soc. Am.*, vol. 45, no. 5, pp.1193-1205, May 1969.

[17] M. Epstein and M. Florentine, "A test of equal loudness ratio hypothesis using cross modality matching functions," *J. Acoust . Soc. Am.*, vol. 99, no. 2, pp. 907-913, August 2005.

[18] L. E. Marks, "Binaural summation of the loudness of pure tones," *J. Acoust . Soc Am.*, vol. 64, no.1, pp. 107-113, July1978.

[19] R. W. L. Kortekaas and A. Kohlrausch, "Psychoacoustical evaluation of PSOLA. II. double-formant stimuli and the role of vocal perturbation," *J. Acoust . Soc. Am.*, vol. 105, no. 1, pp. 522-535, January 1999.

[20] S. Amitay, A. Irwin, D. J. C. Hawkey, J. A. Cowan, and D. R. Moore "A comparison of adaptive procedures for rapid and reliable threshold assessment and training in naïve listeners," *J. Acoust. Soc. Am.*, vol. 119, no. 3, pp. 1616-1624, March 2006.