

# SPEECH PROCESSING FOR HEARING AIDS FOR MODERATE BILATERAL SENSORINEURAL HEARING LOSS

Alice N. Cheeran<sup>1</sup> and Prem C. Pandey<sup>2</sup>

<sup>1</sup>Biomedical Engineering Group

<sup>2</sup>Electrical Engineering Dept.

IIT Bombay, Powai Mumbai-400 076, India.

<alicenc@ee.iitb.ac.in>, <pcpandey@ee.iitb.ac.in>

**ABSTRACT:** Binaural dichotic presentation can be used for reducing the effects of increased temporal and spectral masking associated with sensorineural loss and for improving speech perception by persons with moderate bilateral loss. Speech processing schemes based on adjustable gain filtering, spectral splitting with perceptually balanced comb filters, temporal splitting with trapezoidal fading and inter-aural switching period of 20 and 40 ms, and combined splitting with cyclically swept comb filters with cycle time of 20, 40, and 80 ms were implemented. Experimental evaluation was carried out through listening tests involving recognition of randomly presented words by hearing impaired subjects, for establishing the effect of processing schemes and parameters. Binaural filtering always resulted in improvement. Maximum number of subjects showed large improvement with spectral splitting. Temporal splitting was better with 20 ms switching, while the combined splitting gave best improvements for cycle time of 40 ms.

## 1. INTRODUCTION

The characteristics of sensorineural hearing impairment are frequency dependent shifts in hearing threshold, loudness recruitment, reduced frequency selectivity and increased spectral masking, and reduced temporal resolution and increased temporal masking [1], [2]. With increase in hearing thresholds without corresponding increase in loudness discomfort level, the dynamic range reduces. Increased spectral masking causes smearing of spectral peaks and valleys and reduction of spectral contrast, resulting in degraded reception of consonantal features particularly the place feature. Increased temporal masking leads to increased forward and backward masking of weak acoustic segments by adjacent strong ones. Cues, like voice-onset-time, formant transition, and burst duration, which are important for the identification of consonants, get masked by the vowel segments. Masking takes place primarily at the peripheral auditory level, while integration of information takes place at higher levels in the auditory system. Earlier investigations have shown that processing schemes with binaural dichotic presentation reduced the effect of increased masking for persons with residual bilateral hearing [3]-[6].

An overall improvement of 2 dB in speech-to-noise ratio for dichotic over diotic was reported for splitting with comb filters having eight channel filter bank with constant bandwidths of 700 Hz, designed with complementary interpolated linear

phase FIR filters [3]. Chaudhari and Pandey [4] investigated spectral splitting using a pair of comb filters with complementary magnitude responses based on auditory critical bandwidths [7], with filters designed for sharp inter-band transition. Evaluation of the scheme on subjects with moderate bilateral loss showed significant improvement in recognition scores and perception of consonantal features particularly the place feature.

A pair of comb filters based on 18 critical bands over 5 kHz range was designed for minimum spectral distortion and binaural perceptual balance, as 256-coefficient linear phase FIR filters, using frequency sampling technique [8] applied in an iterative manner treating one or two transition samples as unconstrained and adjusted to obtain the required magnitude response. Listening tests with these filters established that inter-band crossover gain adjusted within 4-6 dB of the pass band gain resulted in perceptual balance and 1 dB ripple in the pass band was found to be acceptable [9]. These filters have transition width of 78 to 117 Hz and stop band attenuation of 30 dB. Listening tests for perceptually balanced filters as compared to comb filters with sharp transition. Cascading each comb filter with a linear phase filter with magnitude response shaped to partly match the audiogram of the test ear, as a partial compensation for frequency dependent hearing threshold shifts, showed further improvement.

A scheme of temporal splitting, in which speech was switched between two ears using trapezoidal fading function with an inter-aural switching period of 20 ms, was investigated [5]. Evaluation on normal subjects with hearing loss simulated at different levels, resulted in improvement of consonantal duration feature, with best improvement for 70 % duty cycle and 2-3 ms transition.

Sensory cells along the cochlear basilar membrane corresponding to alternate auditory bands are always relaxed in spectral splitting. Sensory cells of the two ears are alternately relaxed in temporal splitting. A combined splitting scheme was devised so that all the sensory cells are periodically relaxed from stimulation. The scheme was implemented using a pair of time varying comb filters with pre-calculated set of coefficients, which were selected in steps such that a cyclic sweeping of the pass bands occur. An experimental evaluation conducted on normal subjects with simulated hearing loss with constant sweep cycle of 20 ms provided improvement in the recognition scores and perception of place and duration features [6].

In the present investigation, the four binaural dichotic processing schemes, adjustable gain filtering, spectral splitting

with perceptually balanced comb filters, temporal splitting with trapezoidal fading, and combined splitting with cyclically swept comb filters were implemented. Experimental evaluation was carried out through listening tests involving recognition of randomly presented words by hearing impaired subjects, for establishing the effect of processing schemes and parameters.

## 2. PROCESSING SCHEMES

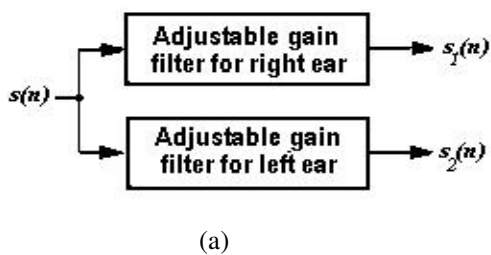
The binaural filtering with adjustable gain response (denoted as AG) permits gain variation within  $\pm 3$  dB to partially compensate for the frequency dependent shifts for each test ear separately. Fig. 1. shows the schematic diagram of the scheme and an example of one subject's pure tone audiogram and the magnitude response of the filter used. The gain variation was restricted because of the limited dynamic range of the hearing impaired subjects. Each filter is a 256-coefficient linear phase filter, designed using frequency sampling technique. These filter pairs were cascaded with the schemes of spectral splitting (SS), temporal splitting (TS), and combined splitting (CS).

Spectral splitting scheme (SS) with perceptually balanced comb filters based on auditory critical bands is shown in Fig. 2. The perceptually balanced comb filters with 9 pass bands each, were designed as 256-coefficient linear phase filter. Fig 3 shows temporal splitting (TS) and the fading functions used for inter-aural switching. Based on earlier results [5], it was decided to use 70 % duty cycle and 3 ms transition duration. Earlier investigation used inter-aural switching period of 20 ms. Here we have used 20 ms (TS-20) and 40 ms (TS-40).

The schematic representation of the combined splitting scheme (CS) with time-varying comb filters is shown in Fig. 4. For an implementation with  $m$  shiftings, each of the time-varying comb filter contained  $m$  perceptually balanced comb filters, which have magnitude responses such that the pass bands of each of these comb filter pairs are shifted in a complementary manner along the frequency axis. Earlier investigation [6] used a sweep cycle of 20 ms, with 2, 4, 8, and 16 shiftings (sweepings). These showed more improvement for 4 and 8 shiftings. Presently combined splitting is considered with sweep cycle of 20, 40, and 80 ms, each with 8 and 16 shiftings. These combinations are denoted as CS-20/8, CS-20/16, CS-40/8, CS-40/16, CS-80/8, CS-80/16.

## 3. EXPERIMENTAL EVALUATION

Experimental evaluation was carried out through listening tests, involving binaural diotic presentation of unprocessed speech (Su)



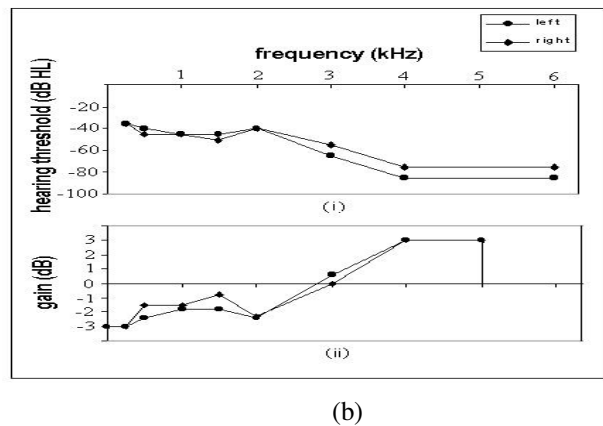
**Fig. 1.** Adjustable gain filter (AG): (a) schematic, (b) pure tone audiogram of a subject and the corresponding filter responses

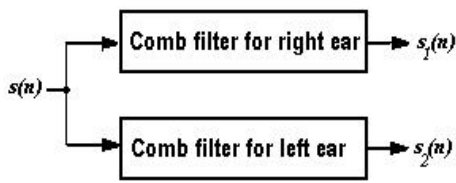
and binaural dichotic presentation of processed speech: binaural filtering with adjustable gain response (AG), filtering cascaded with spectral splitting (SS), filtering cascaded with temporal splitting (TS-20, TS-40), and filtering cascaded with combined splitting (CS-20/8, CS-20/16, CS-40/8, CS-40/16, CS-80/8, CS-80/16).

The listening test material consisted of words presented in a randomized order from a phonetically balanced list of 47 monosyllabic words in Marathi, the first language of the hearing impaired subjects who participated in the tests. This word list is in use at AYJ National Institute for Hearing Handicapped (AYJNIHH), Mumbai, for evaluating speech discrimination by the hearing impaired. The words in the list were recorded at 10 k samples/s, using the line-in of the PC sound card. All the words had approximately the same intensity. The recorded signals were processed off-line for the different combination of processing schemes and parameters. In the listening test set-up, a PC with sound card was used for presentation of the processed signals through the two output channels of the sound card, two audio amplifiers, and a pair of audiometric headphones, to the subject seated in an acoustically isolated room. After each presentation, the subject responded verbally through a microphone, and the response was listened by the experimenter sitting outside and entered on the PC keyboard as right or wrong. In a test, each word was presented 3 times. Each test used a different randomization of words, and it took 15-30 minutes. The listening test program tabulated the responses as well as the response times. Before each test, the subject listened to the words in any order to become familiar with the list and the processed sounds. Thirteen subjects (11 male and 2 female, aged 19-61 years) with mild to severe bilateral sensorineural hearing loss participated in the listening tests.

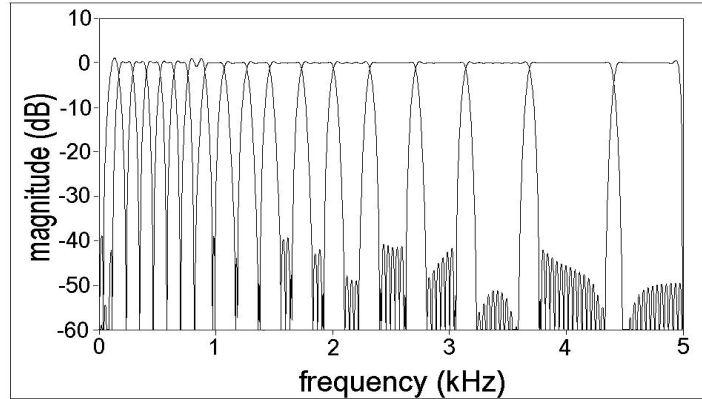
## 4. TEST RESULTS

Fig. 5 shows the recognition scores for all processing conditions tested. The scores for unprocessed speech ranged over 21 - 91 %. The relative percentage improvements because of the various processing schemes and parameters varied across the subjects: 1 - 66 for AG, 6 - 121 for SS, 2 - 110 for TS-20, 0 - 79 for TS-40, -87 - 55 for CS-20/8, -84 - 69 for CS-20/16, -18 - 107 for CS-40/8, and -38 - 138 for CS-40/16, -57 - 110 for CS-80/8, -53 - 117 for CS-80/16. The relative improvements in recognition scores were the least for the combined splitting scheme with 20 ms cycle time (CS-20/8 and CS-20/16). For some subjects, this processing even reduced the scores.



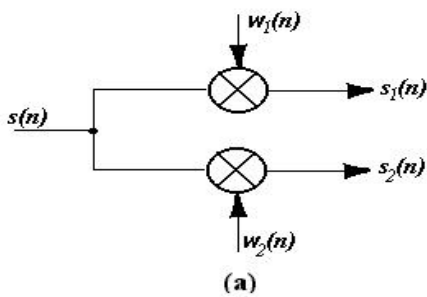


(a)

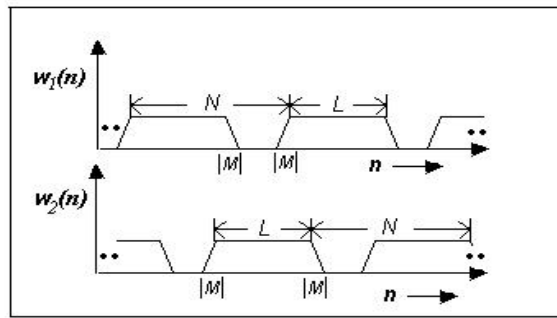


(b)

**Fig. 2.** Spectral splitting (SS): (a) schematic, (b) magnitude response of the comb filter pair, S.R. = 10 k Sa/s, pass band ripple < 1 dB, stop band atten. > 30 dB, transition width = 78 - 117 Hz.

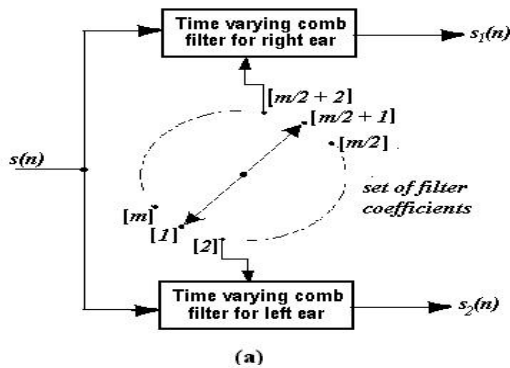


(a)

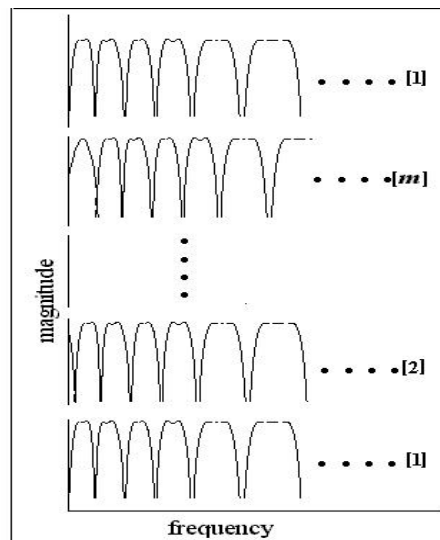


(b)

**Fig. 3.** Temporal splitting (TS): (a) schematic, (b) trapezoidal fading functions with inter-aural switching period of  $N$  samples, duty cycle of  $L/N$ , transition duration of  $M$  samples.



(a)



(b)

**Fig. 4.** Combined splitting (CS): (a) schematic, (b) magnitude response of a time-varying comb filter.

A paired t-test (1-tailed), across subjects, was carried out for testing the statistical significance of the improvements caused by the processing schemes and parameters. Largest relative improvements across the subjects were for SS (31.4%) and TS-20 (30.3%) and both were statistically significant ( $p < 0.03$ ).

In order to study the pattern of improvements, the hearing loss of these subjects was categorized, by their pure tone audiograms. SM and RM, with almost flat loss at low frequencies and increasing at high frequencies with a low slope, had maximum improvement for spectral splitting (15 and 55 %). For

AB and DV, with symmetrical low frequency loss (with maximum loss at 500 Hz), scores degraded for combined splitting for all values of sweep cycle and shiftings. They had maximum improvement with AG (7 and 6 %). Subjects KJ and WK had symmetrical sloping high frequency loss, steeply sloping up to 1 kHz. They showed maximum improvements for SS and TS-20 (22 and 21 % for KJ, 32 and 34 % for WK).

Subjects PK and TT had asymmetrical losses, PK having less loss at mid frequencies and TT having more loss in the frequency range of 3–4 kHz. PK had maximum improvement for SS (6 %), whereas TT had maximum improvement for CS-80/16 (42 %). Subject SS with symmetrical severe mid frequency loss showed improvement for temporal splitting, 32 and 30 % for TS-20 and TS-40 respectively. Subject SK, with severe symmetrical hearing loss and very low recognition score for unprocessed speech (21 %) showed highest improvement with combined splitting CS-40/16 (138%). The subject BS with high frequency loss had more improvement for SS (31%).

### 5. CONCLUSION

Investigations reported here involved implementation and experimental evaluation of speech processing schemes for

improving speech perception by persons with moderate bilateral sensorineural loss. Binaural filtering with gain response adjusted in accordance with the test ear's audiogram improved recognition scores for all the subjects. All the three dichotic schemes had the binaural filtering cascaded with them. Improvements due to processing varied across the subjects. Maximum number of subjects showed improvements with spectral splitting with perceptually balanced comb filters. Between the two intra-aural switching intervals for temporal splitting, improvements were higher for 20 ms. For combined splitting, the sweep cycle of 20 ms gave the poorest results, and best results were obtained for 40 ms. In summary, the dichotic processing schemes implemented in binaural hearing aids should improve speech perception for persons with moderate bilateral sensorineural loss. Further, the processor should permit selection of the dichotic scheme and fine tuning of processing parameters. Combination of these schemes with multiband compression needs to be investigated. Splitting of audio signal may affect sound source localization, particularly for narrow band sounds, and these effects need to be investigated.

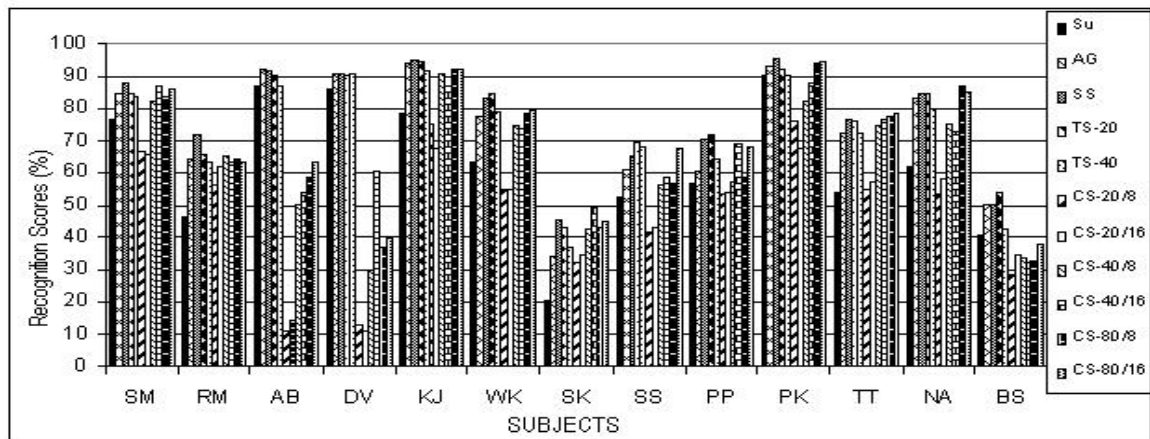


Fig 5. Recognition scores (%) for Unprocessed speech (Su) and for various processing conditions

### REFERENCES

[1] B.C.J. Moore, "Psychoacoustics of cochlear hearing impairment and design of hearing aids," Proc. 16<sup>th</sup> Int. Congress Acoust. (ICA), Seattle, Washington, USA, pp. 2105-2108, 1998.  
 [2] T. Baer and B.C.J. Moore, "Effects of spectral smearing on the intelligibility of sentences in noise," J. Acoust. Soc. Am., vol. 94, pp. 1229-1241, 1993.  
 [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. 38, pp. 75-81, 1993.  
 [4] D.S. Chaudhari and P.C. Pandey, "Dichotic presentation of speech signal with critical band filtering for improving speech perception," Proc. ICASSP'98, Seattle, Washington, USA, paper AE 3.1, 1998.  
 [5] D.S. Jangamashetti and P.C. Pandey, "Inter-aural switching with different fading functions for dichotic presentation to reduce

the effect of temporal masking in sensorineural hearing loss," Proc. SCI'2000, Orlando, Florida, USA, pp.434-439, 2000.  
 [6] A.N. Cheeran, P.C. Pandey, and D.S. Jangamashetti, "Optimal sweep cycle for time-varying comb filters for binaural dichotic presentation to improve speech perception in sensorineural hearing impairment," J. Acoust. Soc. Am., vol. 111(5), pp. 2426, 2001.  
 [7] E. Zwicker, "Subdivision of the audible range into critical bands ( Frequenzgruppen)," J. Acoust. Soc. Am., vol. 33, pp. 248, 1961.  
 [8] L.R. Rabiner, B. Gold, and C.A. McGonegal, "An approach to the approximation problem for nonrecursive digital filters," IEEE Trans. Audio and Electroacoustics, vol. AU-16, pp. 83-105, 1970.  
 [9] A.N. Cheeran and P.C. Pandey, "Dichotic presentation for binaural hearing aids using perceptually balanced critical bandwidth based comb filters," Proc. SCI'2003, Orlando, Florida, USA, 2003.