BINAURAL DICHOTIC PRESENTATION FOR MODERATE BILATERAL SENSORINEURAL HEARING-IMPAIRED

Alice N. Cheeran
Biomedical Engg. Group
IIT Bombay
Powai, Mumbai-400 076, India.
(+91-22) 25764417
alicenc@ee.iitb.ac.in

Prem C. Pandey
Electrical Engg. Dept.
IIT Bombay
Powai, Mumbai-400 076, India.
(+91-22) 25767445
pcpandey@ee.iitb.ac.in

ABSTRACT

For persons with moderate bilateral sensorineural hearing loss, binaural dichotic presentation of speech has helped in improving speech perception by reduction the effect of increased spectral and temporal masking. Speech processing schemes, which split the speech into two for binaural dichotic presentation, are evaluated on normal subjects with simulated hearing loss, and on hearing-impaired subjects with moderate to severe sensorineural hearing loss. Spectral splitting with perceptually balanced comb filter, temporal splitting with trapezoidal fading functions of inter-aural switching periods of 20, 40, and 80 ms, combined splitting with cyclically swept comb filter of sweep cycle of 20, 40, 80 ms and 4, 8 and 16 shiftings is implemented and evaluated by conducting listening tests, which involved recognition of phonetically balanced monosyllables. After the initial round of experiments on seven normal hearing subjects with hearing loss simulated to different levels, processing conditions which provided least improvement were omitted for hearing-impaired subjects. For all the processing conditions considered, the speech was passed through an audiogram dependant filter with gain varying between +3 dB and −3 dB to partially compensate for the frequency dependant shifts in hearing threshold separately for right and left ears, before doing the processing for spectral, temporal, and combined splitting with different conditions, for hearing-impaired subjects. Inter-aural switching period of 20 ms gave better results in temporal splitting, while sweep cycle of 40 ms was preferred in combined splitting. More improvement was for spectral splitting for maximum number of subjects.

Keywords
Sensorineural hearing loss, binaural hearing, dichotic presentations, spectral masking, temporal masking.

Paper Identification Number: CI-3.3

1. INTRODUCTION

Sensorneural hearing impairment is characterized by frequency dependant shifts in hearing thresholds, loudness recruitment, reduced frequency and temporal resolution, and increased spectral and temporal masking [1], [8]. Due to increase in threshold of hearing with no change in loudness discomfort level, the dynamic range reduces, hence speech signals within a much lesser intensity range will only be audible and comfortable Reduction in frequency resolution, acts as though the auditory filters are broader than normal. The charity of speech perceived reduces due to the smearing of the peaks of the speech spectrum. Increased temporal masking leads to increase in 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 following or preceding vowel segment, resulting in degraded speech perception. Masking takes place primarily at the peripheral level, while integration of information takes place at higher levels in the auditory system. Earlier investigations, using schemes that split speech into two for binaural dichotic presentation, has reported to have reduced the effect of increased masking for persons with residual hearing in both the ears [2], [4], [5], [6], [7].

Lunner [7] investigated a scheme of splitting with comb filters having eight channel filter band with constant bandwidth of 700 Hz, designed with complementary interpolated linear phase FIR filters. An improvement of 2 dB in speech-to-noise ratio for dichotic over diotic was reported. In another investigation, Chaudhari and Pandey [2] used a pair of comb filters with complementary magnitude responses based on auditory critical bandwidths, described by Zwicker [11]. The bandwidths were constant at 100 Hz for center frequencies below 500 Hz and were 15-17 % of the center frequency in the range of 1-5 kHz. The filters were designed with the consideration of relatively flat response in the pass band, high attenuation in stop bands and sharp inter-band transition. Evaluation of the scheme on bilateral sensorineural hearing-impaired subjects showed significant improvement in recognition scores and perception of consonantal features particularly the place feature. Further, magnitude response of the comb filters was adjusted to partially match the audiogram of the hearing impaired subjects [3]. Additional improvement was reported but no specific contribution from place feature.

A pair of comb filters based on 18 critical bands over 5 kHz, range was designed as 256-coefficient linear phase FIR filters with sampling rate of 10 k Sa/s, using frequency sampling technique for minimum spectral distortion and perceptual balance For design of such filters, frequency sampling technique was used in an iterative manner, treating one or two transition samples as unconstrained [10]. The unconstrained samples are 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 perceptually acceptable [3]. These filters have transition width of 78 to 117 Hz and stop band attenuation of 30 dB. Listening tests on moderate bilateral sensorineural hearing impaired, showed higher improvement in recognition scores and information transmission of consonantal features for perceptually balanced filters compared to comb filters with sharp transitions. Subsequently, effects of cascading a linear phase filter with magnitude response shaped to partly match the audiogram of the test ear, as a partial compensation for frequency dependent shifts in hearing thresholds was also investigated. Further improvement obtained, maximum for subject having flat losses.

Jangamashetti and Pandey [5], investigated 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. Evaluation on normal subjects with hearing loss simulated at different levels, for different duty cycles and transition durations, resulted in improvement of consonantal duration feature. More improvement was obtained for 70 % duty cycle with 2 and 3 ms transition durations.

In the cochlear basilar membrane, the sensory cells corresponding to alternate bands are always relaxed in spectral splitting, while the sensory cells of the two ears are alternately relaxed in temporal splitting. A scheme of combined splitting was devised, in which all the sensory cells of the basilar membrane are periodically relaxed from stimulation for some time. The scheme is 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], [9].

Presently an overall evaluation of all the three schemes of binaural dichotic presentation, spectral, temporal, and combined splitting with different processing parameters, is carried out by conducting listening tests, on normal hearing subjects with simulated hearing loss. Further listening tests were carried out on mild to moderately severe bilateral sensorineural hearing-impaired persons, for the three processing conditions, after omitting few of the processing parameters for which the improvement on normal subjects, was less. An audiogram dependant filter with maximum gain variation of ± 3 dB, was cascaded with the three splitting schemes to partially compensate for frequency dependant shifts in hearing threshold for each of the test ears. The effect of audiogram dependant filter also was investigated separately for subjects with hearing-impairment. Listening tests involved recognition of randomly presented phonetically balanced monosyllabic words.

2. PROCESSING SCHEMES

Evaluation on normal subjects was done for the schemes of spectral, temporal, and combined splitting, with various processing parameters. Spectral splitting scheme (SS) with perceptually balanced comb filters based on auditory critical bands is shown in Fig. 1. The perceptually balanced comb filters with 9 pass bands were designed as 256-coefficient linear phase FIR filters. Fig 2 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. Constant inter-aural switching period of 20 ms was used in the earlier investigation. Presently, we have used three inter-aural switching periods, 20 ms (TS-20), 40 ms (TS-40) and 80 ms (TS-80).

The schematic representation of the combined splitting scheme (CS) with time-varying comb filters is shown in Fig. 3. 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 4, 8, and 16 shiftings. These combinations are denoted as CS-20/4, CS-20/8, CS-20/16, CS-40/4, CS-40/8, CS-40/16, CS-80/4, CS-80/8, and CS-80/16.

Evaluation was carried on hearing-impaired subjects for the same three binaural processing conditions, but for lesser number of parameters, decided after analyzing the results listening tests on normal subjects. In addition the effect of the audiogram dependant filter alone also was investigated. The binaural filtering with adjustable gain response (denoted as AG) permits gain variation within ±3 dB to partially compensate for the frequency dependent shifts for each test ear separately. Fig. 4. 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 (AG-SS), temporal splitting (AG-TS), and combined splitting (AG-CS). Temporal splitting was conducted for inter-aural switching period of 20 and 40 ms. Combined splitting with 4 shiftings was also omitted.

3. EXPERIMENTAL EVALUATION

The schemes were implemented for off-line processing. Experimental evaluation was carried out through listening tests on normal subjects and hearing-impaired subjects. Different levels of bilateral sensorineural hearing loss was simulated on normal subjects, by adding broad-band Gaussian noise with short-time signal-to-noise ratios (SNR) of 10 ms. In the present evaluation SNRs of ∞, 3, 0, −3, −6, −9 dB were considered. The schemes evaluated on normal subjects were binaural diotic presentation of unprocessed speech (Su) and binaural dichotic presentation of processed speech: with spectral splitting (SS), temporal splitting (TS-20, TS-40, TS-80), and combined splitting (CS-20/4, CS-20/8, CS-20/16, CS-40/4, CS-40/8, CS-40/16, CS-80/4, CS-80/8, CS-80/16)

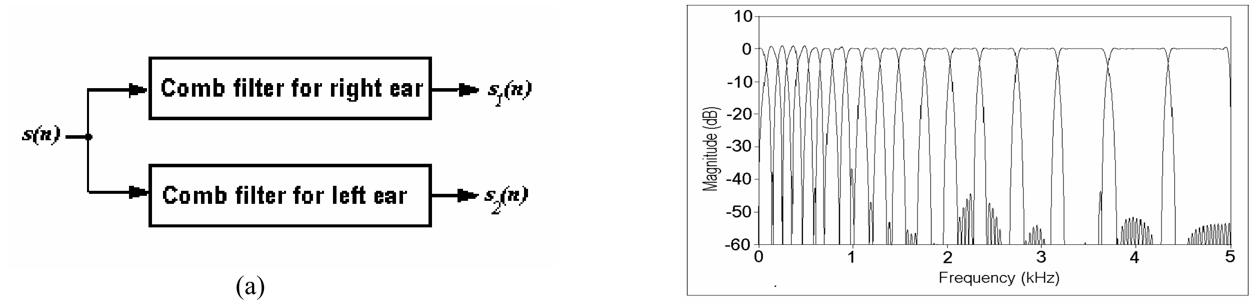

Figure 1: Spectral splitting (SS): (a) Schematic, (b) magnitude response of the comb filter pair, S.R. =10 k Sa/s, pass band ripple < 1 dB, stop band attenuation > 30 dB, = 78 - 117 Hz transition width

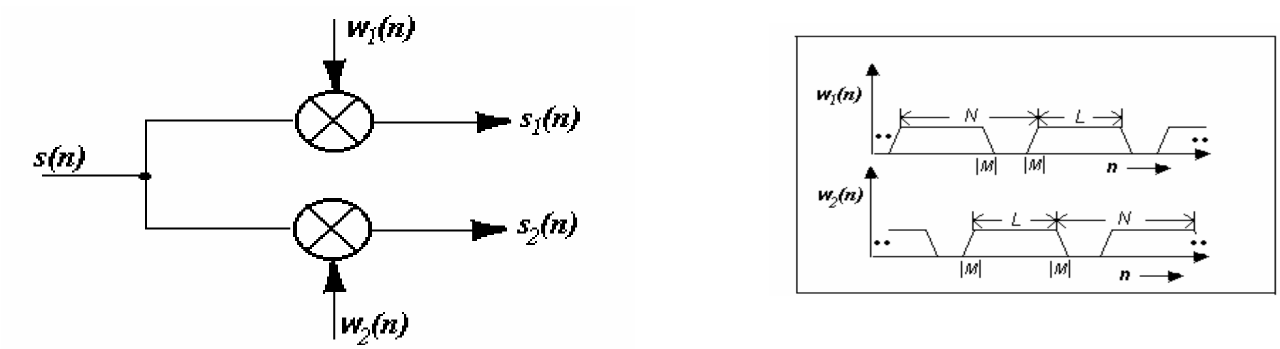

Figure 2: Temporal splitting (TS): (a) schematic (b) trapezoidal fading functions with inter-aural switching period of *N samples,* duty cycle *L/N,* transition *samples* duration of *M samples*

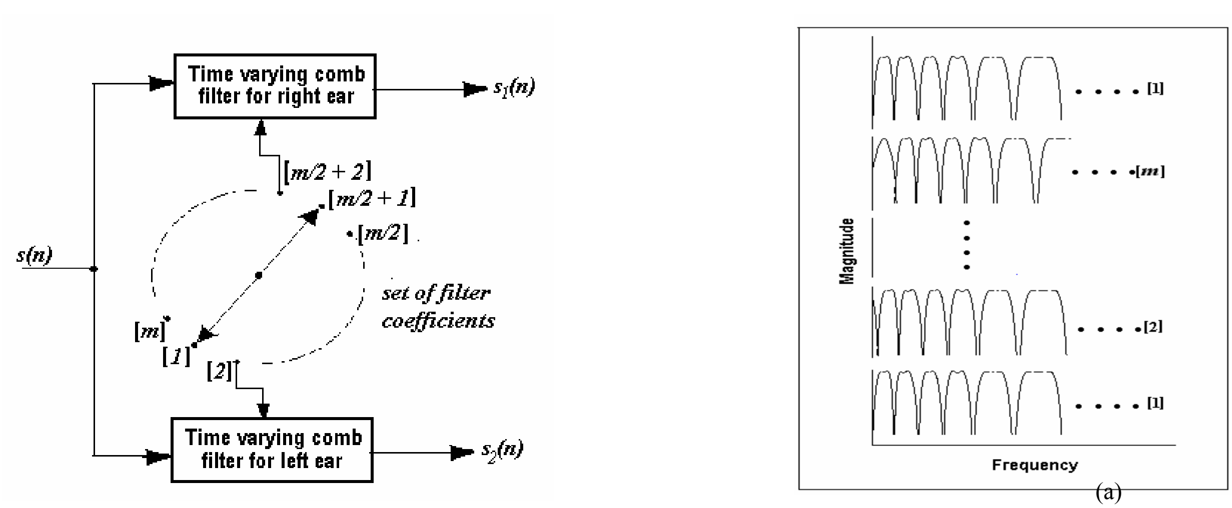

Figure 3: Combined splitting scheme (CS): (a) Schematic and (b) magnitude response of a time-varying comb filter

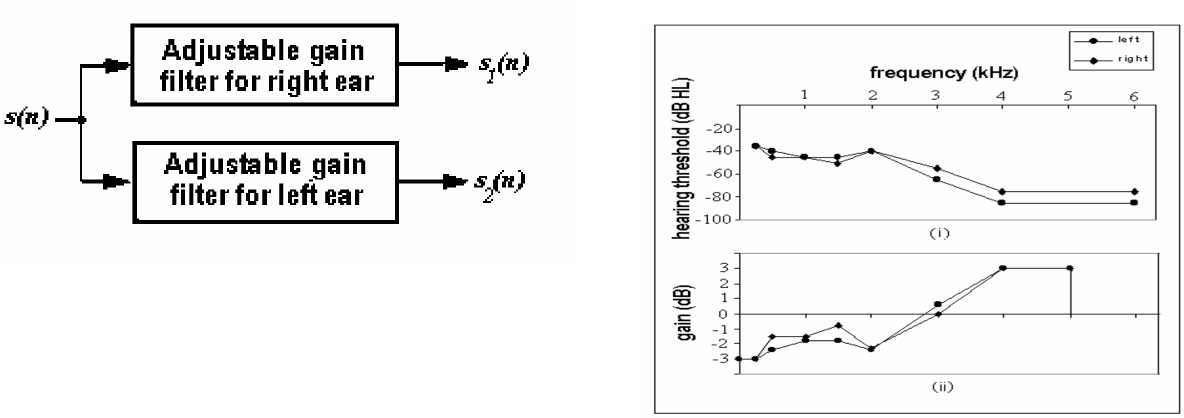

Figure 4: Adjustable gain filter (AG): (a) schematic,
(b) pure tone audiogram of a subject and the corresponding magnitude responses

Experimental evaluation on hearing-impaired subjects was carried out through listening tests, involving binaural diotic presentation of unprocessed speech (Su) and binaural dichotic presentation of processed speech: binaural filtering with adjustable gain response (AG), filtering cascaded with spectral splitting (AG-SS), filtering cascaded with temporal splitting (AG-TS-20, AG-TS-40), and filtering cascaded with combined splitting (AG-CS-20/8, AG-CS-20/16, AG-CS-40/8, AG-CS-40/16, AG-CS-80/8, AG-CS-80/16).

The test material used for listening test consisted of words presented in a randomized order from a phonetically balanced list of 47 monosyllabic words in Marathi, the first language of the 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 8-12 minutes for normal subjects and 15-30 minutes for the hearing-impaired. The listening test program tabulated the responses as well as the response times. Before each test, the subject listened to the words in an order listed, to become familiar with the processed sounds, as many times as required. Seven normal hearing subjects all male (aged 18-30) and thirteen hearing-impaired 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

4.1 Normal Subjects

Fig 5. shows the average recognition scores of unprocessed and processed speech for all processing conditions tested on seven normal hearing subjects for the different SNRs considered. The average recognition score for unprocessed speech reduced from 99.8 % at no noise condition to 23.9 % at −9 dB SNR. At all SNR conditions except at 3 dB, spectral splitting showed highest relative improvement, the scores were 142.8, 93.2, 30.1, 15.7, and 10.9 % for −9 dB, −6 dB, −3 dB, 0 dB, and 3 dB SNRs respectively. For 60 % recognition score. spectral splitting provided an improvement of approximately 5 dB in SNR. Combined splitting with sweep cycle 80 ms with 8 and 16 shiftings resulted in recognition scores close to spectral splitting. The relative improvements for 8 and 16 shiftings were respectively, 130.2 and 118.2 % at −9 dB, 75.4 and 81.0 % at −6 dB, 26.3 and 27.4 % at −3 dB, 12.7 and 13.8 % at 0 dB, and 10.7 and 9.2 % at 3 dB. The SNR improvement for 60 % recognition score was slightly less than 5 dB SNR for combined splitting with sweep cycle of 80 ms for 8 and 16 shiftings. Combined splitting with time varying comb filters are best at 80 ms sweep cycle. The results show very less variation in recognition scores for processing with 8 and 16 shiftings, which is higher than that for 4 shiftings. Under all SNR conditions, the recognition scores for temporal splitting with 80 ms inter-aural switching period was less than that of unprocessed speech. Among the three inter-aural switching periods (20, 40, and 80 ms) chosen, 20 ms showed better results under all SNR conditions. At high SNR conditions the percentage relative improvements in recognition scores for temporal splitting were better compared to low SNR conditions.

4.2 Hearing-Impaired Subjects

Fig. 6 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 AG-SS, 2 - 110 for AG-TS-

20, 0 - 79 for AG-TS-40, -87 - 55 for AG-CS-20/8, -84 - 69 for AG-CS-20/16, -18 - 107 for AG-CS-40/8, and -38 - 138 for AG-CS-40/16, -57 - 110 for AG-CS-80/8, -53 - 117 for AG-CS-80/16. The relative improvements in recognition scores were the least for the combined splitting scheme with 20 ms cycle time (AG-CS-20/8 and AG-CS-20/16). For some subjects, this processing even reduced the scores.

The pattern of improvements was studied, categorizing the hearing loss of these subjects, 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 AG-SS and AG-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 AG-SS (6 %), whereas TT had maximum improvement for AG-CS-80/16 (42 %). Subject SS with symmetrical severe mid frequency loss showed improvement for temporal splitting, 32 and 30 % for AG-TS-20 and AG-TS-20 respectively. Subject SK, with severe symmetrical hearing loss and very low recognition score for unprocessed speech (21 %) showed highest improvement with combined splitting AG-CS-40/16 (138%). The subject BS with high frequency loss had more improvement for AG-SS (31%).

5. CONCLUSION

Speech processing schemes for improving perception of normal hearing subjects in adverse listening conditions and for hearing-impaired subjects with moderate bilateral sensorineural loss, were evaluated. For normal subjects, spectral splitting gave maximum improvement. Temporal splitting was preferred with inter-aural switching 20 ms. Combined splitting scheme gave more improvement for sweep cycle of 80 ms.

For hearing-impaired subjects, 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. 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.

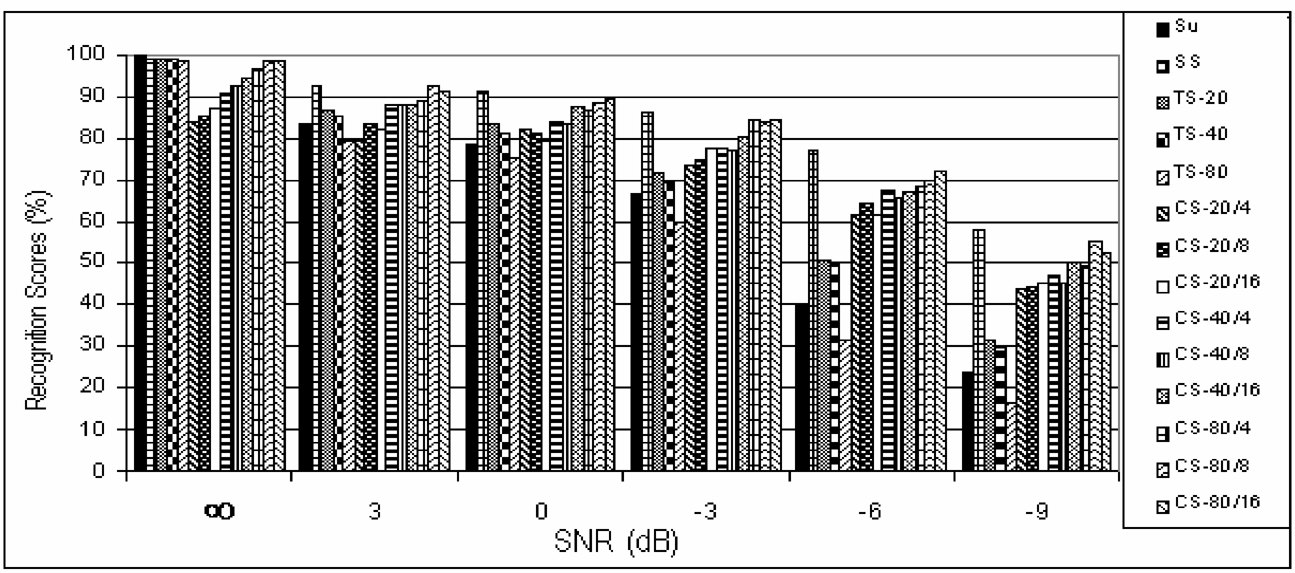

Figure 5: Percentage recognition scores for Unprocessed speech (Su) and processed speech: spectral splitting with perceptually balanced comb filters (SS), temporal splitting with inter-aural switching period of 20, 40, and 80 ms (TS-20, TS-40, and TS-80), combined splitting with sweep cycle of 20, 40, and 80 ms, with 4, 8, and 16 shiftings (CS-20/4, CS-20/8, CS-20/16, CS-40/4, CS-40/8, CS-40/16, CS-80/4, CS-80/8, CS-80/16) for listening tests conducted on normal subjects for simulated hearing loss of ∞, 3, 0, -3, -6, -9 dB -SNRs, averaged over seven.

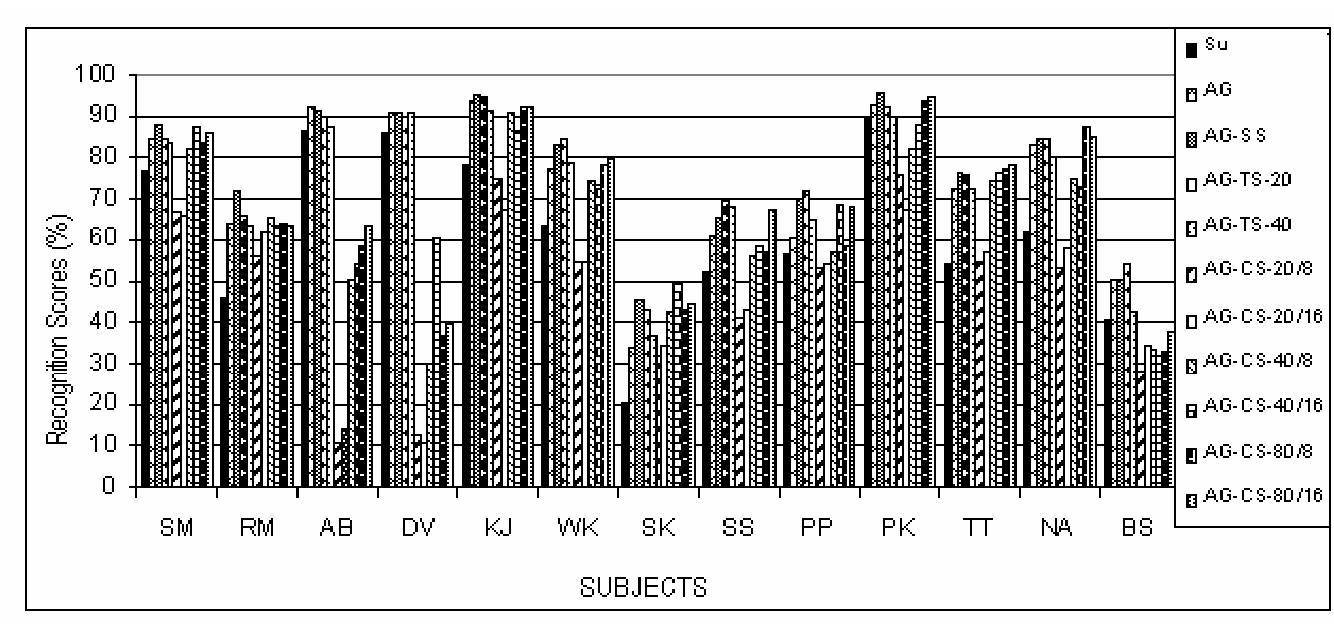

Figure 6: Percentage recognition scores for Unprocessed speech (Su) and processed speech: filtering with adjustable gain response (AG), filtering cascaded with spectral splitting with perceptually balanced comb filters (AG-SS), filtering cascaded with temporal splitting with inter-aural switching period of 20, 40, and 80 ms (AG-TS-20, AG-TS-40, and AG-TS-80), filtering cascaded with combined splitting with sweep cycle of 20, 40, and 80 ms, with 4, 8, and 16 shiftings (AG-CS-20/4, AG-CS-20/8, AG-CS-20/16, AG-CS-40/4, AG-CS-40/8, AG-CS-40/16, AG-CS-80/4, AG-CS-80/8, AG-CS-80/16) for listening tests conducted on normal subjects, averaged over seven.

6. REFERENCES

[1] Baer T., and Moore, B.C.J. "Effects of spectral smearing on the intelligibility of sentences in noise," J. Acoust. Soc. Am., vol. 94, pp. 1229-1241, 1993.

[2] Chaudhari, D.S., and Pandey, P. C. "Dichotic presentation of speech signal with critical band filtering for improving speech perception," Proc. ICASSP'98, Seattle, Washington, USA, paper AE 3.1, 1998.

[3] Chaudhari, D.S., and Pandey, P. C. "Real time speech processing for dichotic presentation for binaural hearing aids," Proc. National Conf. Biomedical engineering, Manipal, Karnataka, India, paper II-45, 1998.

[4] Cheeran, A.N. and Pandey, P.C. "Dichotic presentation for binaural hearing aids using perceptually balanced critical bandwidth based comb filters," Proc. SCI'2003, Orlando, Florida, USA, 2003

[5] Jangamashetti, D.S., and Pandey, P.C. "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] Jangamashetti, D.S., Pandey, P.C., and Cheeran, A.N. "Time varying comb filters to reduce effect of spectral and temporal masking in sensorineural hearing impairment," Proc. Int. Conf. on Biomedical Engineering (ICMBE), Bangalore, India, pp.258-263, 2001.

[7] Lunner, T., Arlinger, S., and Hellgren, J. "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.

[8] Moore, B.C.J. "Psychoacoustics of cochlear hearing impairment and design of hearing aids," Proc. 16th Int. Congress Acoust. (ICA), Seattle, Washington, USA, pp. 2105-2108, 1998.

[9] Pandey, P.C., Jangamashetti, and D.S., Cheeran, A.N., "Binaural dichotic presentation Binaural dichotic presentation to reduce the effect of increased temporal and spectral masking in sensorineural hearing impairment," J. Acoust. Soc. Am., vol. 110(5), pp. 2705, 2001.

[10] Rabiner, L.R., Gold, B. and McGonegal, C.A. "An approach to the approximation problem for nonrecursive digital filters," IEEE Trans. Audio and Electroacoustics, vol. AU-16, pp. 83-105, 1970.

[11] Zwicker, E. "Subdivision of the audible range into critical bands (Frequenzgruppen)," J. Acoust. Soc. Am., vol. 33, pp. 248, 1961.