Wide-Band Compensation of Presbycusis

Chiara Grasso[1], Davide Quaglia[1], Laura Farinetti[1], Giotto Fiorio[3], Juan Carlos De Martin[2]

[1]Dipartimento di Automatica e Informatica/[2]IEIIT-CNR

Politecnico di Torino, C.so Duca degli Abruzzi 24, I-10129 Torino, Italy

[3]IEI-CNR, Via Moruzzi 1, I-56124 , Pisa, Italy

E-mail: [grasso|davide.quaglia|farinetti|demartin]@polito.it, fiorio@iei.pi.cnr.it

ABSTRACT

This paper deals with a wide-band audio filtering technique for the compensation of presbycusis. Instead of restoring frequencies only up to 4 kHz (obtaining telephone–like quality), we propose to preserve frequency components up to 8 kHz, however possibly attenuated by unrecoverable hearing loss. Wide–band speech, in fact, carries important perceptual information that enhances both speech naturalness and intelligibility. A filtering technique that compensates hearing loss up to a given cut-off frequency is presented. A formal Diagnostic Rhyme Test was performed to test the intelligibility of speech processed using the proposed technique versus regular telephone–bandwidth speech. The results show a significant intelligibility gain, proving the potential of wide-band hearing aids for the compensation of presbycusis.

KEY WORDS

Wide-Band Filtering, Presbycusis, Intelligibility, Diagnostic Rhyme Test

1 Introduction

One of the most common form of hearing loss is presbycusis, which is normally related to the body aging process and to long-term exposure to sound energy. The loss associated with presbycusis is usually greater for high-frequency sounds. This contributes to difficulty in hearing and understanding speech, especially for female voices and in background noise. In particular, high-frequency sounds such as "s" and "th" are difficult to tell apart.

The design of an effective hearing aid is a challenging problem because of the nature of the human ear. The most important change that occurs in the impaired ear is the loss of sensitivity to sound energy that can be measured through an audiogram. In addition, other issues have to be taken into account, e.g., the reduced frequency resolution, the dead regions of the cochlea, the recruitment, and the masking of some spectral components [1]. This first stage of our work was mainly devoted to investigate the problem of sensitivity loss and to build up the testing process for the Italian language.

The evolution of hearing aids in recent years has been characterized by a miniaturization process which has delivered very small wearable devices. Presbycusis can be compensated by hearing aids that, among other advanced functions, amplify high-frequency sounds. Usually the frequency range of these devices is limited to telephone-bandwidth [2, 3] to avoid high-power consumption that would result in short battery life. Fig. 1 shows the typical maximum gain curve of a commercial aid, characterized by a rapid gain decrease between 3 and 7 kHz.

Although the speech energy resides mainly in the range from 100 Hz to about 5-6 kHz [4], signal components at higher frequencies contribute to improve the intelligibility and the naturalness of speech. In particular, impulsive or noisy consonant sounds can occur in the range of 2 kHz to about 9 kHz and may be as much as 27 dB lower in amplitude than vowels.

French and Steinberg [5] investigated the intelligibility of low-pass filtered speech. Their test material consisted of meaningless monosyllables of the consonant-vowel-consonant type. Intelligibility was measured as the percentage of syllables in which all the component sounds were perceived correctly. They reported an average intelligibility of 90% for nonsense syllables low-pass filtered at 3.3 kHz. Pollack [6] found the same results using lists of monosyllabic words.

This work aims at testing the intelligibility improvement that can be obtained using a wide-band presbycusis compensation filter. The proposed filter compensates for the hearing loss up to 4 kHz, while its gain remains constant thereafter up to 8 kHz. In this context, consonants are more vulnerable to signal impoverishment since they have high frequency components; therefore the Diagnostic Rhyme Test (DRT) was used, because it tests the intelligibility for different classes of consonant phonemes.

The paper is organized as follows. Section 2 describes the filter used in the tests. Section 3 introduces the Diagnostic Rhyme Test and its adaptation for the Italian language. Experimental method is described in Section 4 and results are reported in Section 5. Finally, some conclusions are drawn in Section 6.

2 The Proposed Filter

The filter used in this work was designed on the basis of the hearing loss of a specific subject affected by presbycusis, as measured by his audiogram. The lower curve in Fig. 2 represents the sound attenuation as a function of the frequency
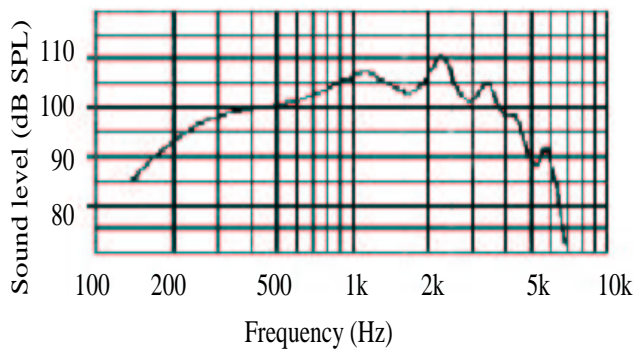
Figure 1. Typical gain curve of commercial hearing aids.

Figure 3. Resulting loss curve after compensation.

for this individual; it shows a 10 dB loss at 1 kHz, 20 dB loss at 2 kHz, and 40 dB loss at 4 kHz.

The proposed filter has finite impulse response and linear phase. Its gain perfectly compensates for hearing loss up to a "critical frequency", equal in this case to 4 kHz; the gain remains constant thereafter. The upper curve in Fig. 2 shows the proposed filter designed for this loss curve. Fig. 3 shows the resulting loss curve after compensation. The hearing loss is perfectly compensated up to the frequency of 4 kHz; at higher frequencies, the gain of the compensation filter remains constant, while the loss due to presbycusis strongly increases.
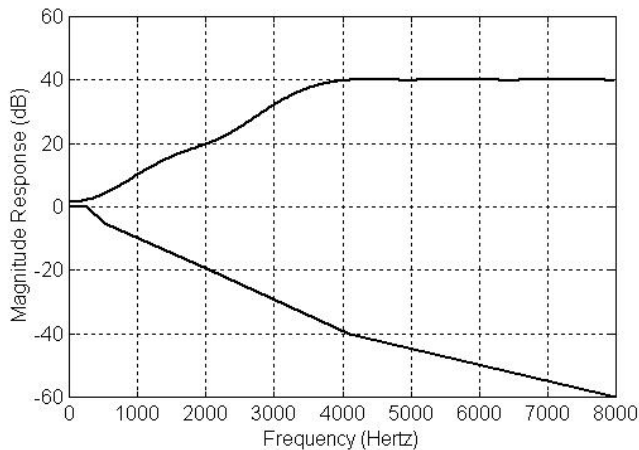

Figure 2. Gain curve of a presbycusic subject (bottom) and the corresponding linear–phase compensation filter (top).

3 The Diagnostic Rhyme Test

3.1 Overview

Statistical intelligibility measurements have been extensively used for the assessment of speech communication systems [7]. A quantitat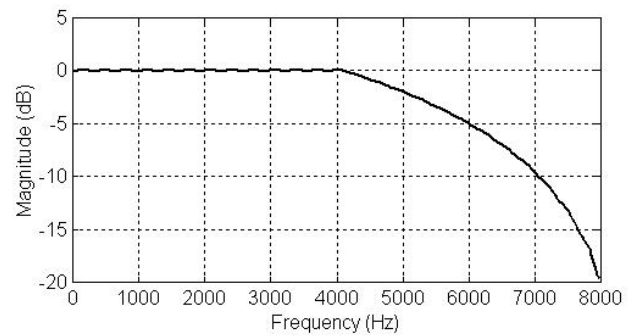ive measure of the intelligibility of speech can be obtained by counting the number of discrete speech units correctly recognized by a listener. Three classes of speech units are normally used: *a)* single syllables consisting of meaningless combinations of speech sounds; *b)* meaningful isolated words; and *c)* meaningful sentences. The main drawback of the use of nonsense syllables is that it requires trained testing crew. On the other hand, psychological factors may affect the results using actual words or, even more, sentences. A more comprehensive study of these issues can be found in [8].

The Diagnostic Rhyme Test (DRT) [9] is more properly described in terms of a set of principles for item construction and selection than in terms of a specific corpus of test materials. In the DRT, one hundred and ninety two words are arranged in ninety-six rhyming pairs, the initial consonants of which differ by a single distinctive feature (e.g., "veal–feel"). Listeners are shown a word pair, then asked to identify which word was presented by the talker. The DRT is based on a number of distinctive features of speech, and its results reveal errors in discrimination of initial consonant sounds. The DRT classifies the stimulus words according to six different features, that for the English language are *voicing*, *nasality*, *sustention*, *sibilation*, *graveness*, and *compactness*.

3.2 Italian Language DRT

Since this work was conducted in Italy and testing session had to be done using the subjects' native language, an adaptation of DRT to the Italian language was necessary. The application of the DRT method to the Italian language requires a suitable classification of speech features and the preparation of proper lists of rhymed words whose first consonant differs for a unique feature. On the basis of a taxonomy of phonemes for the Italian language, in 1986 Bonaventura, Paoloni, Canavesio and Usai [10] developed an adaptation of DRT to the Italian language consisting of one hundred and five word pairs. This set covers six features named *nasalità*, *continuità*, *stridulità*, *coronalità*, *anteriorità*, and *sonorità*. For each feature all the vocal contexts are examined.

4 Experimental Method

4.1 Simulation of the Hearing Loss

The improvement of the wide-band filter has been evaluated considering intelligibility with respect either to a general model of current commercial aids and to perfect loss compensation.

Subjects involved in this test were required to have no hearing impairment and this particular type of hearing loss has been simulated by an appropriate filtering of the input speech stimuli.

Combining the two contributions given by the presbycusis loss and the wide-band compensation filter, shown in Fig. 2, the resulting filter actually tested on normal hearing subjects is shown in Fig. 3. It is worth to note that it is this final curve that has been tested, since the proposed filter is supposed to perfectly adapt its gain to the hearing loss up to a "critical frequency", equal to 4 kHz, and to remain constant thereafter, regardless of the specific case of loss.

The actual commercial aids are assumed to perfectly compensate hearing loss up to a frequency of 3,400 Hz, with a rapid gain decrease thereafter [2, 3]. Therefore they have been simulated by processing the input speech stimuli with a telephone-bandwidth filter.

In order to establish a reference case to evaluate the previous two test conditions, the simulation of perfect compensation has been tested. Since the subjects were required to have no hearing impairment, this was obtained by presenting the input stimuli in their original form.

4.2 Test Procedure

The test was conducted using a software program running on a personal computer equipped with high-quality headphones. The test was divided into a screening session, used to evaluate the subject's pure-tone hearing thresholds, and a longer DRT intelligibility session.

In the screening session, after a short explanation, seven sequences of ten pure-tones are presented to the subject. Sequences are relative to the frequencies of 125, 250, 500, 1,000, 2,000, 4,000 and 8,000 Hz. Each sequence consists of ten tones, which are 100 ms in length and whose intensity decreases in 10 steps of 5 dB each. Each sequence is repeated twice and then a dialog window asks to the subject how many tones he/she believes to have heard. This preliminary session aims at validating the assumption that the subject has no hearing impairment.

The intelligibility session is divided into three separated parts, each lasting about ten minutes. Between one part and the following, the subject is asked by the program to take a rest. In the first part telephone-bandwidth speech stimuli are presented; in the second part stimuli filtered with the proposed wide-band technique, and in the last part the speech stimuli are not filtered. Each part presents the 210 words of the DRT set, each of them preceded by the carrier sentence "Now select ___" (in Italian, "Adesso seleziona ___"). In this way the listener is prepared for the presentation of the test item and variability in the scores due to inattention is reduced [11]. After each speech stimulus, a dialog window is displayed, containing two buttons labeled with the words between which the subject has to choose the one he/she believes to have heard.

5 Results

The test was performed on twenty-two volunteers, fifteen males and seven females. They are native Italian speakers, aging between 20 and 40 years, of normal hearing. For each test case and for each consonant feature the number of errors was recorded. The number of errors has been normalized using the formula suggested by Voiers [9] and reported in (1).

$$S = 100\frac{R - W}{T} \qquad (1)$$

In (1) S is the corrected normalized score, R is the number of correct responses, W is the number of incorrect responses and T is the total number of items involved (in this experiment, T is equal to 210). This correction takes into account the possibility that the subject chose the right answer only by guessing. With this formula, every wrong answer is counted twice, since it assumes that an other error got undetected.

Results have been grouped and classified; Fig. 4 shows the distribution of the number of subjects that achieved a given score. The three filtering approaches show the same distribution among the different intervals.
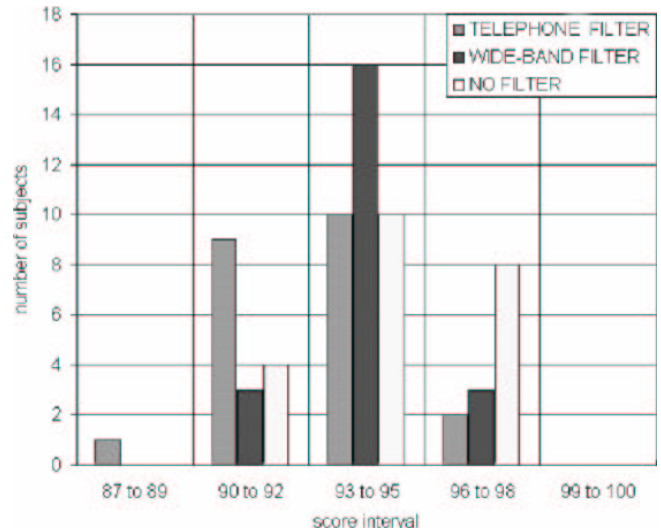

Figure 4. Distribution of the number of individuals achieving a given score.

A statistical analysis has been performed on the results. With a 95% confidence level, the average score relative to the telephone filter belongs to the interval 93.09% ±

0.95, i.e., from 92.14% to 94.04%. Using the same confidence level, the average score relative to the wide-band filter belongs to the interval 95.00% ± 0.74 (Interval 2). With unfiltered speech stimuli the resulting interval is 95.40% ± 0.82 (Interval 3). Fig. 5 shows the 95% confidence intervals for the three different cases. Interval 1 and 2 are statistically disjoint, i.e., with a 95% confidence the two cases are statistically different, with the wide-band filter outperforming the telephone filter. Interval 2 and 3, instead, are statistically non separable; this means that the proposed technique is approximately as intelligible as the original input speech.
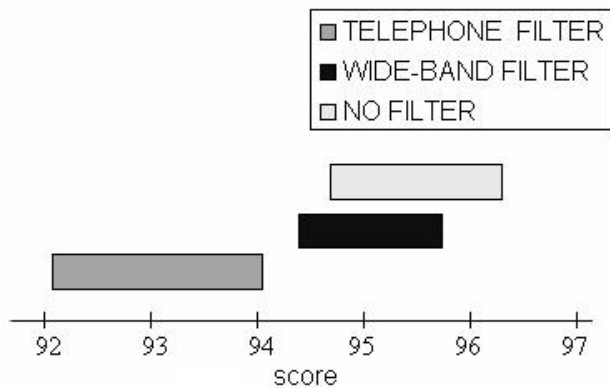

Figure 5. Statistical analysis of the resulting scores: the 95% confidence intervals relative to the three Test Cases.

Fig. 6 shows the distribution of errors for each consonant feature. It is evident from the histogram that the number of errors that are related to the features *stridulità* and *sonorità* is significantly lower with the wide-band filter than with the telephone filter. This behavior can be explained considering that these features are characterized by high-frequency energy components, which are strongly attenuated by the telephone filter. As a verification, fig. 7 and fig. 8 report three spectrograms relative to the words "silo" and "filo", respectively (they are one of the word pairs for the feature *stridulità*). For each word, the upper spectrogram is relative to it in its original form, the middle when it is filtered with the wide-band filter and the lower when it is filtered in the telephone bandwidth. While in the case of the wide-band filter the two spectrograms are still distinguishable, the two words, filtered in the telephone bandwidth, have, in practice, the same spectrum.

For the features *coronalità* and *nasalità* the number of errors reported in Fig. 6 is practically independent of the filtering technique.

6   Conclusion

This work presented the perceptual results of a filtering technique to compensate for hearing loss due to presbycusis. Traditional hearing aids restore frequencies only up to 4 kHz, obtaining telephone–like quality and losing high-frequency components that may improve both speech intel-
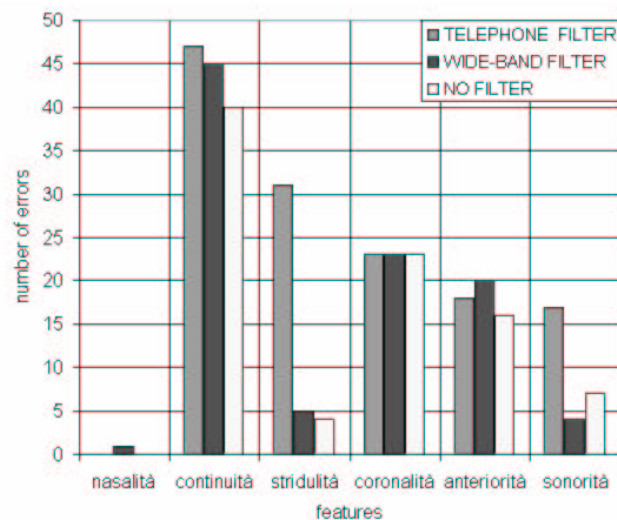

Figure 6. Distribution of scores for each consonant feature.

ligibility and naturalness. We studied the improvement in intelligibility when exact compensation is extended up to 4 kHz and frequency components up to 8 kHz, albeit attenuated, are preserved. The proposed wide-band filtering approach was designed on the basis of an actual loss curve of a specific subject. A statistical assessment of intelligibility was performed using the Diagnostic Rhyme Test on a group of twenty-two individuals. The statistical analysis of the results demonstrates that the proposed wide-band filtering technique outperforms the narrow-band filtering, proving the potential of wide-band hearing aids for the treatment of presbycusis.

References

[1] A. Maynard Engebretson, Benefits of digital hearing aids, *IEEE Engineering in Medicine and Biology Magazine*, 13(2), 1994, 238–248.

[2] A. Ben Hamida, Implication of new technologies in deafness healthcare: deafness rehabilitation using prospective design of hearing aid systems, in *Proceedings of the IEEE International Symposium on Technology and Society*, 2000, 85–90.

[3] G. Bertini, G. Fiorio, and M. Marani, Limits of hearing aids in the high frequency range, *Proceedings of the 17th International Congress on Acoustics*, 2001.

[4] E. Zwicker and H. Fastl, *Psycho-acoustics: Facts and Models* (Springer, 1999).

[5] N.R. French and J.C. Steinberg, Factors governing the intelligibility of speech sounds, *Journal of the Acoustical Society of America*, 19(1), 1947, 90–119.
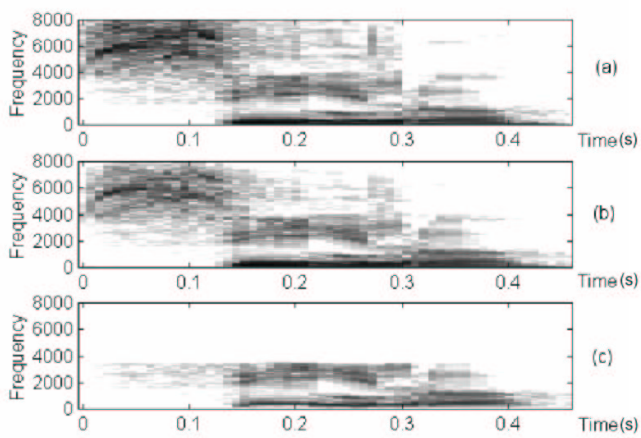
Figure 7. Spectrograms of the word "silo": (a) no filter, (b) wide–band filter, and (c) telephone filter. Spectrograms were computed using a sliding Hanning window of 16 ms.
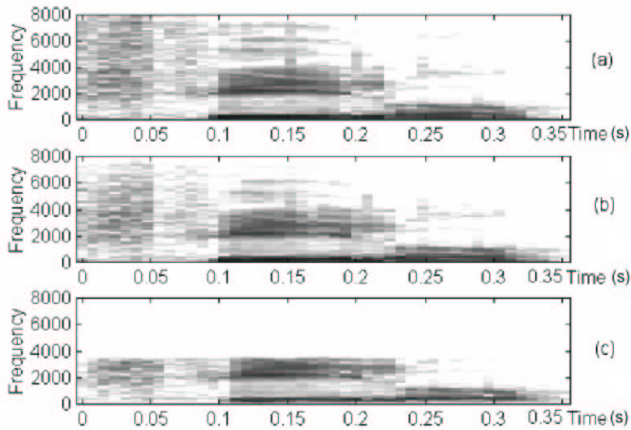

Figure 8. Spectrograms of the word "filo": (a) no filter, (b) wide–band filter, and (c) telephone filter. Spectrograms were computed using a sliding Hanning window of 16 ms.

[6] I. Pollack, Effects of high pass and low pass filtering on the intelligibility of speech in noise, *Journal of the Acoustical Society of America*, 20(3), 1948, 259–266.

[7] Mones E. Hawley, Ed., *Speech Intelligibility and Speaker Recognition* (Stroudsburg, PA: Dowden Hutchinson and Ross, 1977).

[8] H. Fletcher, *Speech and Hearing in Communication* (Huntington, NY: Krieger, 1953).

[9] W.D. Voiers, Diagnostic evaluation of speech intelligibility, in Mones E. Hawley (Ed.), *Speech intelligibility and speaker recognition* (Stroudsburg, PA: Dowden Hutchinson and Ross, 1977).

[10] P. Bonaventura, A. Paoloni, F. Canavesio, and P. Usai, Realizzazione di un test diagnostico di intellegibilità per la lingua italiana, in *Convegno Associazione Italiana Acustica*, 1986.

[11] J.P. Egan, Articulation testing methods, *Laryngoscope*, 58(9), 1948, 955–991.