FORMANT-BASED PROCESSING FOR HEARING AIDS

P.J. Blamey*, G.J. Dooley*, P.M. Seligman+, J.I. Alcantara*, and E.S. Gerin*

* Human Communication Research Centre
University of Melbourne

+ Cochlear Proprietary Limited

ABSTRACT - A body-worn hearing aid has been developed with the ability to estimate formant frequencies and amplitudes in real time. These parameters can be used to enhance the output signal by "sharpening" the formant peaks, by "mapping" the amplitudes of the formants onto the available dynamic range of hearing at each frequency, or by resynthesizing a speech signal that is suited to the listener's hearing characteristics. Initial evaluations have indicated small improvements in speech perception for three groups of subjects: users of a combined cochlear implant and speech processing hearing aid, normally hearing listeners in background noise, and a hearing aid user with a severe hearing loss.

SOME PROBLEMS WITH HEARING AIDS

Most conventional hearing aids amplify sounds with a fixed linear gain function that varies across frequencies to compensate for the hearing loss. Commonly, the maximum power output is also limited to avoid uncomfortably loud sounds. These aids make sounds louder, but do not overcome all of the problems that are associated with hearing-impairment. For example, recruitment is a narrowing of the dynamic range of hearing caused by threshold levels being raised more than discomfort levels for sound. This effect is a consequence of the loss of outer hair cells that produce the very sensitive and highly tuned cochlear response to pure tones at low levels (Patuzzi, 1990). When the vulnerable outer hair cells are lost through exposure to loud sound, disease, or other trauma, the sensitivity is lost, but responses to louder sounds are relatively unaffected. Fine frequency resolution and selectivity can also be reduced by loss of the outer hair cells (Evans, 1975). These distortions of the hearing sensations operate to reduce the intelligibility of speech even when it is amplified to a comfortable level in quiet conditions. Background noise is also a major problem for hearing aid users who are usually affected much more than normally hearing listeners. In the last two decades, it has been proposed that hearing aids should be designed to compensate for the effects of recruitment, poor frequency resolution, and background noise as well as providing amplification (e.g. Villchur, 1973).

Multiband compression has been a common method used to compensate for recruitment. The amplitude levels for fixed frequency bands are compressed nonlinearly to match the dynamic range of hearing (e.g. Lippmann, Braida & Durlach, 1981). At most, this method has led to modest improvements in speech intelligibility compared with linear gain. The limiting factors appear to be distortions in the processed signals and reduced loudness contrasts between different temporal and spectral components of the speech. Spectral enhancement is a procedure that increases the amplitude differences between peaks and valleys of a spectrum with the aim of compensating for reduced frequency selectivity (Simpson, Moore & Glasberg, 1990). A narrowband enhancement algorithm can increase the relative amplitude of the harmonic components of the voice. A wideband procedure increases the relative amplitude of the formants. These procedures have produced small improvements in speech intelligibility for hearing impaired listeners in laboratory studies. Directional microphones have been used in hearing aids for some time to reduce the effects of background noise. Processing to implement adaptive beamforming microphones also shows promise (e.g. Peterson, Durlach, Rabinowitz & Zurek, 1987). Adaptive filtering based on long-term noise spectra is less effective in improving intelligibility, but can decrease the annoying effects of some types of noise (e.g. Levitt, Neuman, Mills & Schwander, 1986).

In addition to the fundamental problems outlined above, there are practical difficulties in determining and fitting ideal linear gain functions to the characteristics of an individual user's hearing loss.

Skinner (1988) has compared rules for determining ideal gain, and digital hearing aids are likely to be successful in fitting these ideal functions closely in practical situations (e.g. Levitt *et al*, 1986).

A SPEECH PROCESSING HEARING AID

Although hearing aids that use the algorithms mentioned above are sometimes called "speech processing" hearing aids, they do not usually make much use of acoustic properties that are specific to speech signals. This paper provides a brief description and a summary of initial results for a hearing aid that makes use of formant frequencies and amplitudes, estimated in real time, with the aim of enhancing acoustic information that is known to be important in the perception of speech.

The device is based on the cochlear implant speech processor developed by the University of Melbourne and Cochlear Pty Ltd. A brief description and further references are included in Blamey & Tartter (this volume). The implant processor has been modified by the addition of a programmable filter circuit to produce an acoustic output signal. The result is a very flexible hearing aid with access to the fundamental frequency, formant frequencies and amplitudes measured by the speech processor. The device is in use as a "bimodal" aid by implant users who have residual hearing in the non-implanted ear, and as a speech processing hearing aid by people with severe to profound hearing losses that are not easily fitted with conventional hearing aids.

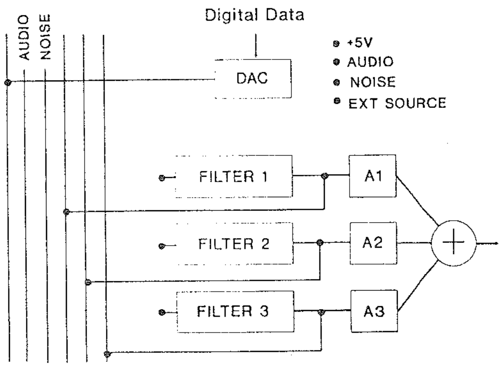

Figure 1. Block diagram of the acoustic processing chip

The low-powered CMOS filter circuit was implemented on a single chip application-specific integrated circuit (ASIC). The basic structure is shown in Figure 1. It contains three switched-capacitor biquad filters (Gregorian & Temes, 1986) that can be configured as low-pass, band-pass or high-pass in parallel or serial arrangements. The centre frequency (50 Hz to 25 kHz), bandwidth (0.01 to 4 octave), and gain (-31.5 to +31.5 dB in steps of 0.5 dB) of each filter are dynamically programmable. The input to each filter can be selected from the speech signal, a digital-to-analog converter, a white noise source, an external signal source, or the output of either of the other two filters. The outputs of the filters are summed and amplified on the chip to drive a standard hearing aid output transducer. The configuration of the chip and the filter parameters are controlled by loading digital codes into latches on the chip from the cochlear implant processor. Thus the device can be used with a fixed configuration or can be controlled dynamically in response to the changing speech parameters estimated by the implant speech processor.

MODES OF OPERATION

The device is particularly flexible and can be used in a number of different modes. These modes allow accurate frequency response setting, spectral shaping to emphasize dynamically changing formants, accurate loudness control of dynamic signals, resynthesis of speech signals, or variations and combinations of these schemes. A few of these modes have been chosen for initial evaluation, and are discussed below.

Frequency Response Tailoring (FRT)

In this mode of operation, the filters are programmed to produce a fixed frequency response as close as possible to the prescribed real ear gain for the user's hearing loss across the frequency range from 250 Hz to 6 kHz (Byrne & Dillon, 1986). An iterative gradient search method is used to find the least squares fit of the complex transfer function of the filters to the desired filter gain. For most hearing losses it has been possible to implement a frequency response within 1 dB of the prescribed gain. The result is similar to the fitting of a conventional hearing aid but with some advantages: a) The fitting of the ideal gain function is more accurate than is possible with the adjustments to high, middle and low frequency gain and slope that are available on some hearing aids. b) The adjustments can be made quickly without sending the hearing aid to the manufacturer. c) The audiologist can alter the ideal gain function if required to suit the user. d) Different gain prescription rules can be implemented and compared quickly by the user. e) The hearing levels of the user can be measured using the hearing aid itself under computer control. This removes the need for correction factors to model the effect of an ear mould since the effect is included in the measurement.

Peak Sharpening (PS)

In a more complex mode of operation called "peak sharpening", the filters are programmed dynamically to emphasize the first (F1) and second (F2) formants of the speech signal. The speech signal is passed through two bandpass filters connected in parallel. The centre frequencies of the filters are dynamically adjusted to track the F1 and F2 estimates produced by the cochlear implant speech processor. The resultant signals are amplified to the appropriate level specified by the Byrne & Dillon (1986) ideal gain rule for the user's hearing loss. Filter parameters are updated at a rate equal to the fundamental frequency of the voice during voiced sounds, and at a faster pseudo-random rate for unvoiced sounds. Theoretically, PS could enhance the perception of formant information by hearing-impaired users with poor frequency selectivity, and improve signal-to-noise ratios in moderate amounts of background noise. The filter bandwidths are important parameters for PS because too broad a setting will have little effect on the spectrum while too narrow a setting will result in uncontrolled amplitude variations as harmonic peaks move in and out of the filter pass-bands. The effect of filter bandwidth on speech perception has been evaluated in two of the preliminary studies described below.

Loudness Mapping (LM)

A refinement of the PS mode of operation can be used to control loudness and compensate for recruitment effects. Measurements of the loudness of narrowband noise at different frequencies and levels can be used to produce loudness growth curves and iso-loudness contours for individual hearing aid users. Comparison with loudness data for normally hearing subjects will then indicate the gain required to produce normal loudness as a function of the input signal's frequency and intensity. In the LM mode, the required gain will be applied to two filters that dynamically track F1 and F2 as in the PS mode. This approach differs from nonlinear multiband compression because the gain functions applied are linear at all times and do not distort the spectral information in the signal.

Resynthesis (RS)

The RS mode uses the F0, F1, and F2 parameters from the implant speech processor to produce a completely synthetic speech signal. The possible advantages of this approach include complete control over the intensity and frequency characteristics of the output signal, a simplified speech spectrum which may be beneficial in cases of poor spectral resolution and selectivity, and the ability to encode selected information in ways that may be more salient to the listener. The latter possibilities include transposition of information to lower frequencies for listeners who have no usable hearing at high frequencies, and the provision of F0 and amplitude information with the aim of improving lipreading as in the SiVo aid (Faulkner, Ball, Rosen *et al*, 1992).

PRELIMINARY EVALUATIONS

Bimodal evaluation with cochlear implant users

The combined implant/hearing aid processor has been evaluated with a group of five implant users including four with a severe hearing loss in the unimplanted ear (Dooley, Blamey, Seligman *et al*, in press) and one with a profound hearing loss in the unimplanted ear. The subjects were tested with a benchtop version of the processor using the FRT and PS strategies combined with the normal MPEAK implant coding scheme. The control condition was their normal implant speech processor and their normal hearing aid as independent devices. The results shown in Figure 2 were obtained for a recorded open-set monosyllabic word test (Boothroyd, 1968) presented from a single speaker at 70 dB SPL in a sound-attenuating room and scored phonemically. With only one exception, scores with the bimodal processor were at least as high as with the subjects' own independent devices, despite their lack of practice with the benchtop processor. The average score in the best bimodal condition was significantly higher than in the independent devices condition [Mean difference=10.9%, $t(4)=2.38$, $p<0.05$, Student t-test for paired samples]. The filter bandwidth used for the PS condition in this study was 1/3 octave. The improvement with the bimodal aid is possibly attributable to the accurate fitting of the ideal gain, and the use of a single microphone for the binaural signal. Patients using independent devices sometimes report that the two sounds do not fuse together well and can be confusing. It is likely that the very different processing in the independent devices accentuates any differences in the frequency responses and amplitude envelopes of the two signals. These differences are minimised in the bimodal device. Further bimodal developments will use data from binaural pitch and loudness studies to match the hearing sensations from the electric and acoustic signals in more detail.

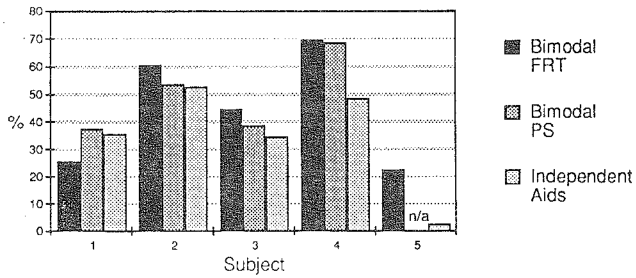

Figure 2. AB Word Test scores for users of a combined implant and hearing aid speech processor.

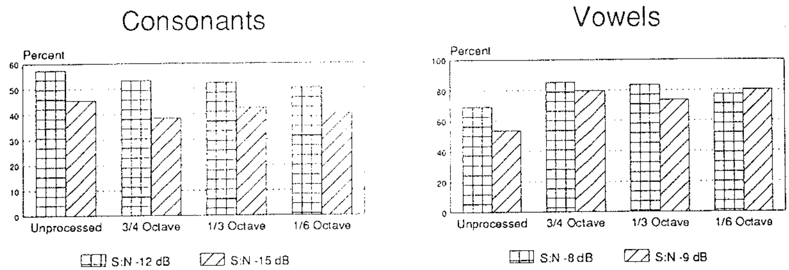

Figure 3. Vowel and consonant scores for normally hearing listeners using the PS mode in background noise.

Peak sharpening in noise with normally hearing listeners

The aim of this study was to investigate the effect of PS on the perception of speech in noise and the effect of changing the bandwidth of the filters. Four normally hearing adults were evaluated with a set of 11 vowels in /hVd/ context and 12 consonants in /aCa/ context at two signal-to-noise (S:N) ratios and three bandwidths (3/4, 1/3, and 1/6 octave). The S:N ratios were chosen separately for the vowels and consonants to reduce the recognition scores to about 50% in the unprocessed condition. The noise used was multitalker babble. The experiment was concerned with perceptual effects rather than the robustness of the formant extraction process in noise, so the formant parameters used were derived from the speech signal before the noise was added. The signal plus noise was then processed with the wearable processor in PS mode. As shown in Figure 3, the processing produced a small improvement in the perception of vowels but not consonants. Scores were slightly higher for wide bandwidths compared to narrow bandwidths.

Peak sharpening in quiet with a hearing-impaired listener

Initial results on a sentence test in quiet (see Figure 4) suggest that the 1/2 to 3/4 octave bandwidth range may be most suitable for a hearing-impaired listener, in good agreement with the data for normally-hearing listeners. This result also suggests that the 1/3 octave bandwidth may not have been optimal in the bimodal study described above.

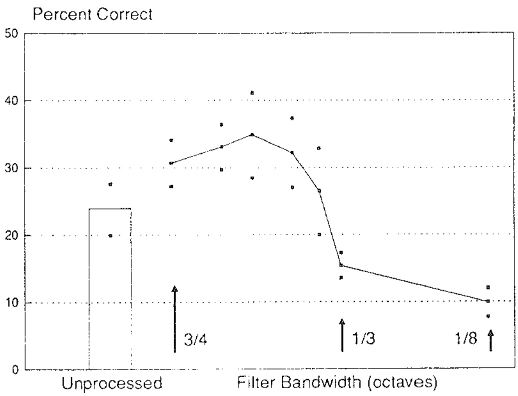

Figure 4. CUNY Sentence Test scores for one hearing-impaired listener using the PS mode.

CONCLUSIONS

The aid described above is capable of implementing a wide variety of speech processing strategies. It is likely that users with different hearing losses will be most appropriately aided with different strategies. For example, it has been suggested that people with profound hearing impairment may benefit from hearing aids that measure and encode specific speech features (e.g. Levitt, 1986; Boothroyd, 1990; Faulkner *et al*, 1992). Several resynthesis strategies of this type are currently being investigated with the formant-based hearing aid. Users with more hearing, whose main problems may be poor frequency selectivity or recruitment may be more appropriately fitted with spectral enhancement strategies such as peak sharpening or loudness mapping. All users of the aid should benefit from the accurate control of the gain to achieve the desired frequency response.

The studies carried out so far indicate that modest improvements in speech perception in quiet and in noise may be achieved using formant-based processing for hearing aids. Much more thorough evaluation and optimisation of the processor is needed to validate this conclusion and to ensure that the maximum improvement is attained. It is clear that the processor has the flexibility and power required to carry out these investigations.

REFERENCES

Blamey P.J. & Tartter V.C. (1992) *Fricative perception by cochlear implant users,* Proc. Fourth Aust. Conf. Speech Science & Technology, SST-92, Brisbane (this volume).

Boothroyd, A. (1968) *Developments in speech audiometry,* Sound 2, 3-10.

Boothroyd, A. (1990) *Signal processing for the profoundly deaf,* Acta Otolaryngol. (Stockh) Suppl. 469, 166-171.

Byrne, D. & Dillon, H. (1986) *The National Acoustic Laboratories' new procedure for selecting the gain and frequency response of a hearing aid,* Ear & Hearing 7, 257-265.

Dooley, G.J., Blamey, P.J., Seligman, P.M., Alcantara, J.I., Clark, G.M., Shallop, J., Arndt, A., Heller, J. & Menapace, C.M. (in press) *Combined electrical and acoustical stimulation using a bimodal prosthesis,* Arch. Otolaryngol.

Evans, E.F. (1975) *The sharpening of cochlear frequency selectivity in normal and abnormal cochleae,* Audiology 14, 419-442.

Faulkner, A., Ball, V., Rosen, S., Moore, B.C.J. & Fourcin, A. (1992) *Speech pattern hearing aids for the profoundly hearing-impaired: Speech perception and auditory abilities,* J. Acoust. Soc. Am. 91, 2136-2155.

Gregorian, R. & Temes, G.C. (1986) *Analog MOS integrated circuits for signal processing,* New York: John Wiley, p280-296.

Levitt, H. (1986) *Hearing impairment and sensory aids: A tutorial review,* J. Rehab. Res. Dev. 23(1), xiii-xviii.

Levitt, H., Neuman, A., Mills, R. & Schwander, T. (1986) *A digital master hearing aid,* J. Rehab. Res. Dev. 23(1), 79-87.

Lippman, R.P., Braida, L.D. & Durlach, N.I. (1981) *Study of multichannel amplitude compression and linear amplification for persons with sensorineural hearing loss,* J.Acoust. Soc. Am. 69, 524-534.

Patuzzi, R. (1990) *Mammalian cochlear function,* in Rowe, M. & Aitkin, L. (Eds) *Information processing in mammalian auditory and tactile systems,* Wiley-Liss: New York, 45-59.

Peterson, P.M., Durlach, N.I., Rabinowitz, W.M. & Zurek, P.M. (1987) *Multimicrophone adaptive beamforming for interference reduction in hearing aids,* J. Rehab. Res. Dev. 24(4), 103-110.

Simpson, A.M., Moore, B.C.J. & Glasberg, B.R. (1990) *Spectral enhancement to improve the intelligibility of speech in noise for hearing-impaired listeners,* Acta Otolaryngol. (Stockh.) Suppl. 469, 101-107.

Skinner, M.W. (1988) *Hearing aid evaluation,* Prentice Hall: New Jersey.

Villchur, E. (1973) *Signal processing to improve speech intelligibility in perceptive deafness,* J. Acoust. Soc. Am. 53, 1646-1657.