# PATTERN ELEMENT HEARING AIDS AND SPEECH ASSESSMENT AND TRAINING

Adrian FOURCIN and Evelyn ABBERTON

## Summary

This paper has been prepared for a meeting (in Beijing 9-10 IX 1996) organised jointly by the Chinese Academy of Sciences and the Department of Phonetics and Linguistics at UCL. The meeting is concerned with new technology hearing aids, with special reference to phonetically important speech dimensions and bio-active implants.

The particular aim here is to give examples of speech element processing with special reference to perceptual and productive assessment. The same basic principles apply to assessment and training for both prosthesis and patient, however, and they can be briefly defined by the following summary points concerning the present and potential advantages of using essential speech dimensions:-

- clarity of perception and relevance to communication

- targeted signal processing

- complementary assessment of prosthesis and patient

- simply structured adult rehabilitation and training

- matching acquisition to normal stages of development

## Background

The pattern element hearing aid is intrinsically designed to achieve three principal objectives. First, to target the extraction from speech signals of components which are directly related to important phonetic dimensions. Second, to employ signal processing methods which utilise these dimensions so as to get a degree of immunity from the effects of external noise. Third, to provide a rational perceptual basis for matching the operation of the aid to the individual listener's auditory, and language, abilities.

Ordinarily, basic speech dimensions such as loudness, voice, frication, timbre and nasality are readily perceived as distinct entities. However, at the level of the acoustic signal these essential speech dimensional components are not clearly defined physically but embedded relatively inaccessibly in the time course and frequency spectrum of the signal. This has made their robust application difficult. Current commercially successful cochlear stimulation prostheses operate, in consequence, on the basis of the whole signal - by providing the user with a simple running approximation to the speech spectral envelope, sampled rather coarsely in frequency and relatively rapidly in time. The approach is rather that of the original articulation index, the principles of which were first worked on by Collard and Fletcher (1929). In essence the whole signal is also presented by conventional acoustic hearing prostheses although they may employ quite complex signal processing. Methods of receptive and productive assessment used in conjunction with these prostheses tend to mirror their

techniques of processing and VCV, whole word, sentence and discourse level appraisal and teaching techniques are the most important ones applied.

The present work is concerned with a parallel approach to that normally followed. Although phonetic dimensions are at the heart of the speech signal they are not ordinarly used for structured training and assessment in work for totally deaf and hearing impaired people. The employment of speech pattern elements which map on to particular phonetic dimensions requires analytic techniques which are additional to those ordinarily employed both in assessment and in training. Once these means of analysis are available, the design and functioning of the prostheses themselves are also likely to be  profoundly modified. This is because the same principles of training and assessment can be used both to design and set up the hearing aid and to train and assess the human user.

In associated work Faulkner (1992) describes some of the receptive results obtained by deaf users when listening and lip-reading with a particular pattern element aid, SiVo (Fourcin 1990),which gives loudness, pitch and frication information; and Wei discusses the speech pattern element training used to define the signal analysis characteristics of the SiVo signal  processor (1992).  In this SiVo work there is an intrinsic complementarity between the training and assessment of the hearing aid and the training and assessment of the human user.  The following examples are intended briefly  to  illustrate  these  processes.   Special  reference  is  made  to  the  physical correlates of the auditory dimensions of: loudness; voice pitch; nasality and frication.

## Speech Pattern Element Display and Analysis
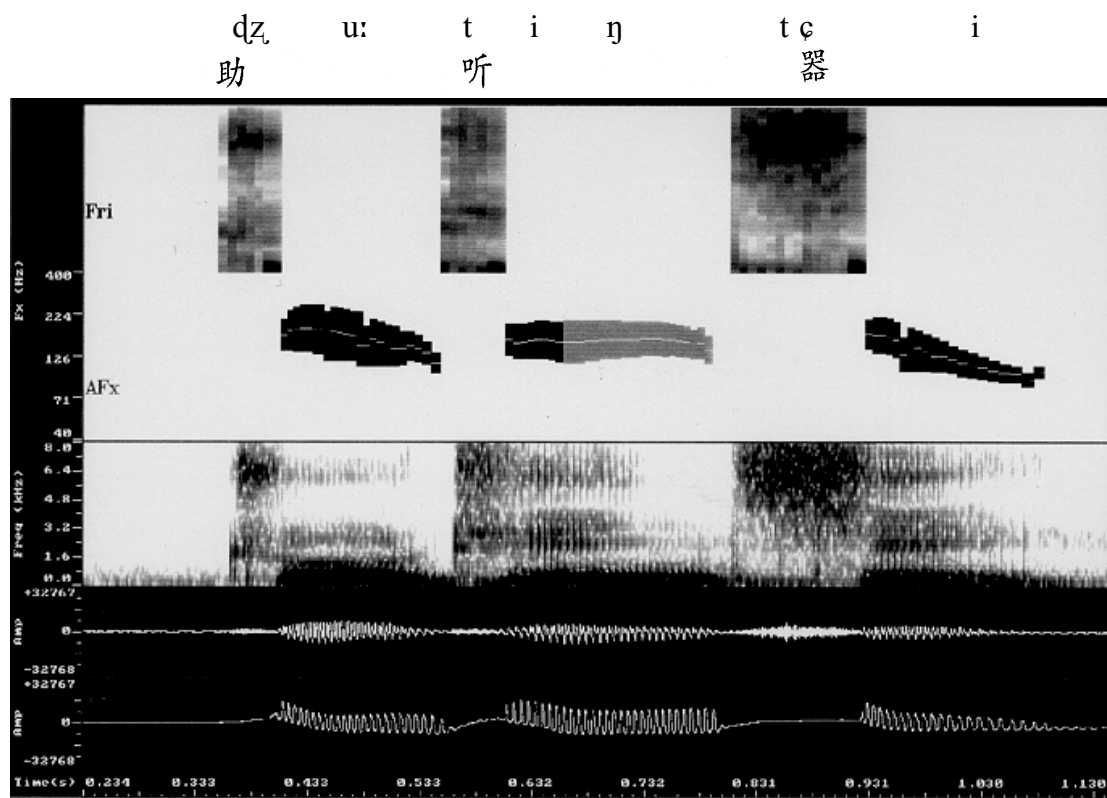


The figure is for the utterance:-        **Zhuting qi**  (hearing aids)

*Adrian Fourcin & Evelyn Abberton*

represented in two ways as the result of real-time analysis. The stylised spectrum above (both spoken and computer processed by Xinghui Hu using Laryngograph Ltd sensors) corresponds to the pattern elements of frication, loudness, voice pitch and nasality. Below, there is a standard wide-band frequency-time spectrogram. In the pattern spectrum, voice pitch is shown as a vertically width modulated trace where the width at any instant corresponds to the amplitude of the speech signal and the shading intensity relates to the presence or absence of nasality. Frication is shown above this voice trace by rectangular blocks within each of which there is an indication of the frequency spread of the frication energy contained in the speech signal as a function of time. Whereas the pattern spectrum is quite simple, the complete wide band spectrogram below is complex. The physical correlates of voice pitch, fundamental frequency and period are not very easy to identify visually in the spectrogram and nasality is even more obscure. Frication, on the other hand, is quite distinct but even here the pattern spectrum is much clearer.

Our work is beginning to make more use of this type of data clarification (Toffin et al 1995), and the current signal analytic algorithm development has been based on the use of this type of speech element data. Its availability in real time as an interactive display is, in a complementary fashion, of direct utility for user familiarisation and training (Abberton 1975). The analyses shown make it feasible to examine analytically aspects of speech production which are directly related to these receptive dimensions since their phonetic definition is both in regard to auditory ability and speaking skills The following discussion is intended to give examples at the levels of voice, nasality and frication.

**Speech Output and Laryngeal Control with SiVo and with a Conventional Aid**
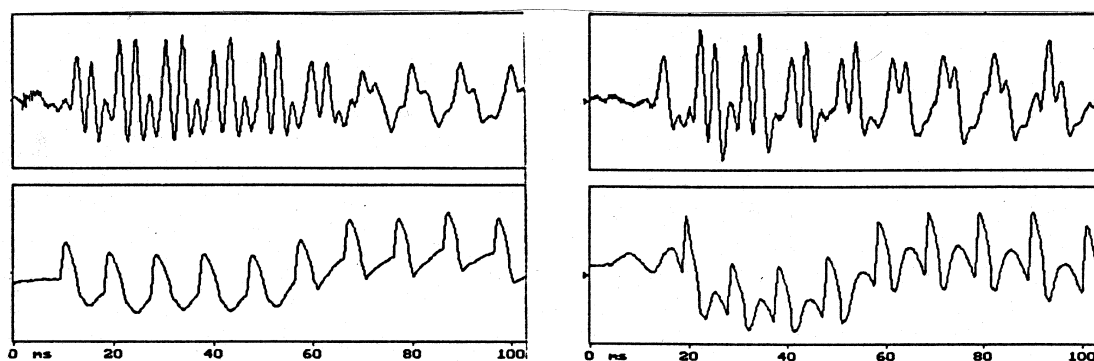


Speech Output and Lx control with SiVo  Speech Output and Lx control with CHA

The pairs of waveforms above show the plots for the acoustic signal of speech, Sp, above, and the laryngograph signal, Lx, below.

On the left the pair of waveforms is for the condition where the profoundly deaf user (MR) has been monitoring his speech using a SiVo aid programmed for his comfortable auditory levels, and on the right where he has been using the familiar

conventional amplifying hearing aid which he uses daily. This hearing aid user (one of the patients in Dr Elisabeth Fresnel's clinic in France), has made many recordings of this passage during the course of the work and since it is an overlearned situation he has got into the habit of using the same patterns of prosodic control. This makes it possible to understand what would otherwise be an astonishing difference between the two sets of waveforms above. The conventional aid, which has been used in the recording is not able to provide MR with an adequate sensation of voice pitch and in order partially to overcome this lack of sensory feedback, he has a tendency to produce a breathy excitation spectrum - which has proportionately more energy at the low end of the spectrum. The Sp waveforms show this fairly clearly, but there is an evident anomaly in the Lx waveform. The use here of the SiVo aid to monitor his speech activity has not only enabled MR to control the broad levels of his phonation but also to produce normal vocal fold closure sequences throughout his read passage. The sample shown is at the beginning of the word "Séguin". (The Sp & Lx waveforms are as initially recorded, without time alignment. Note the correspondence in the underlying Lx baseline in the two sets of recordings). In fact, throughout the whole of this recording the larynx excitation traces are normal when the SiVo hearing aid has been used for self monitoring by MR and abnormal when he monitors his speech with the conventional hearing aid. This change in speaking control is dependent on the type of hearing aid used and occurs necessarily within the same clinical session.

**Speech Production with and without auditory feedback using SiVo**
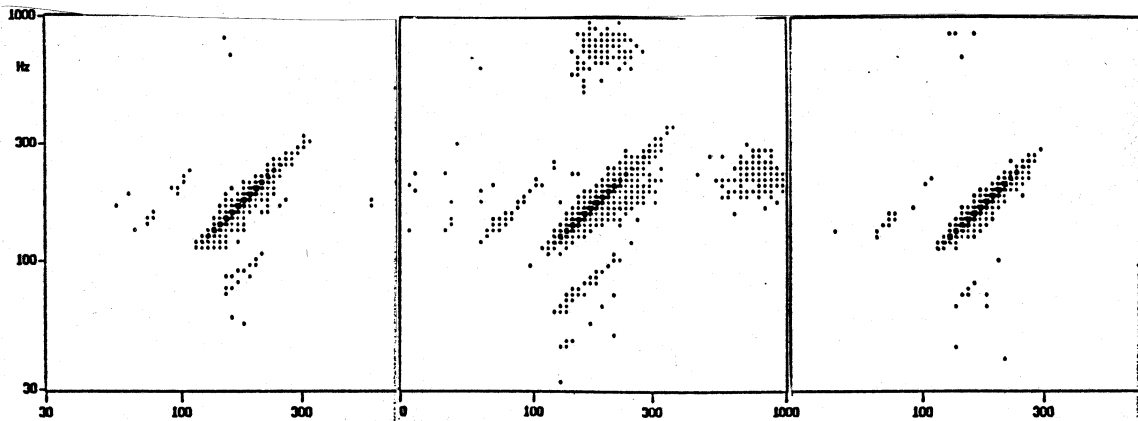


A complementary aspect of self monitoring comes from the comparison of the effects of the use of a prosthesis with no provision at all. The set of Fx crossplots shown above has been made from a sequence of recordings in which the (UK) user of a SiVo aid has, from left to right: first used the aid to assist in reading a standard passage; then after an hour read again with no aid; and then immediately again used the SiVo

aid for self monitoring.  At the beginning and end of the entire session, using the SiVo for monitoring, the distribution means and irregularity measures stay the same. In the absence of auditory control more variability between productions is typically found and there is also more variability within a given spoken sequence.  Mean & modal Fx values differ, and irregularity alters in both nature and magnitude.   When a conventional aid is used the errors in speech production tend to be the same since they are based on an incorrect auditory representation.

## Distributions of Pattern Element Activity

### Speech Pattern Elements
#### s29s1.pat



### Speech Pattern Elements
#### s22s4.pat
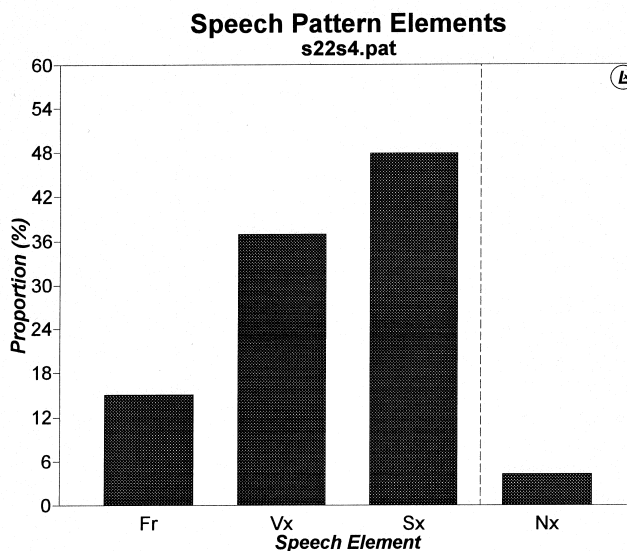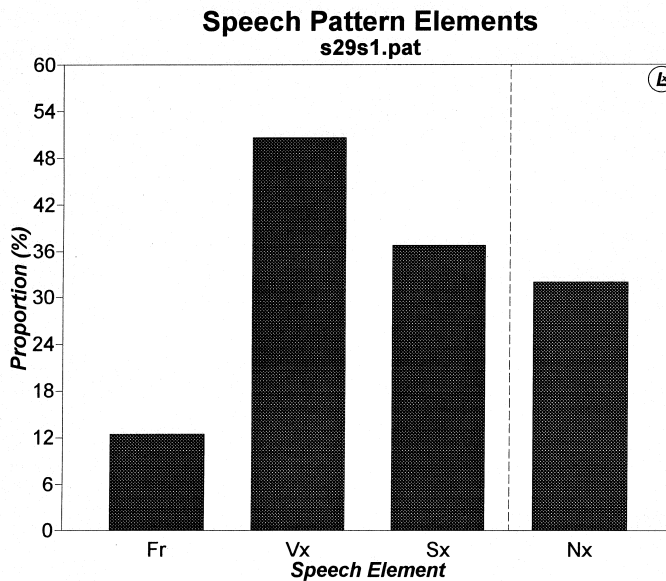


Important aspects of both speech perception and speech production relate directly to the basic phonetic dimensions which are used in the interactive display shown on page

*Pattern element hearing aids and speech assessment training*

36. It would be an advantage if the speech production characteristics resulting from the use of a hearing aid user could be shown simply along these lines. The two histogram plots above give examples of the application of this approach. The horizontal axis for each covers the dimensions of: frication; voicing; silence and nasality. In the first plot, the speaker (although normal) has inappropriate nasalisation. This is clearly shown by the comparison with the right hand plot where the same speaker has spoken normally. These analyses are based on the use of nasal and laryngograph sensors together with a normal microphone. In the SiVo both laryngograph and frication analyses are replicated by the use of trained neural net processors, where the training uses targets derived from these sensor sources of information. The analyses shown in the figure above are of use in the evaluation of a hearing aid user's ability to benefit from the aid and any pattern element based training. They also enable the prosthesis itself to be evaluated by the use of synchronous reference speech and sensor signals.

## Discussion

This brief overview has given particular examples of improved productive control resulting directly from the use of speech pattern element hearing aids. Here, improved perceptual ability has been based on the use of voice targets. This is of especial importance for Chinese and tone language environments. The hearing aid design which has made these results possible is based on the use of the same principles of analysis and training which are used for patient assessment and training. This symmetry of approach is likely to become more widespread in future generations of hearing aids because it directly links human need to signal processing technique. A structured signal analytic method of this "human" type is not, however, restricted to a minimal set of features. In principle as close an approximation to a complete speech signal as is desired may be obtained. The structured use of the resulting family of speech pattern elements can be used, as here, for analytic work. It can also be employed as a match to the normal processes of speech and language acquisition by the hearing impaired child.

## Acknowledgements

## References

Collard, J. (1929) A theoretical study of the articulation and intelligibility of a telephone circuit *Electl Commun*.,7,168-186

Fletcher, H. & Steinberg, J. C. (1929) Articulation testing methods'.*Bell Syst Tech. J.*, 8, 806-854

Faulkner, A., Ball, V., Rosen, S. R. Moore, B.C.J., Fourcin, A. (1992) Speech Pattern Hearing Aids for the profoundly hearing impaired: Speech perception and auditory abilities *JASA* 91, 4pt1, 2136-2155

Wei, J., Faulkner, A., Fourcin,A. (1992) Speech pattern processing for Chinese listeners with profound hearing loss *Proceedings ICA 14 Beijing*

Fourcin, A. (1990)  Prospects for speech pattern element aids *Acta Otolaryngol (Stockh) Suppl.* 469, pp 257-267

Toffin,C, et al, (1995) Voice production as a function of auditory perception with a speech pattern element hearing aid  *Proceedings ICPHS  Stockholm* 3, pp 206-209

Abberton, E. & Fourcin, A. (1975) Visual Feedback and the Acquisition of Intonation in Lenneberg, E & E (eds) *Foundations of Language Development* Volume 2 chapter 1 pp 157-165, Academic Press NY