VOICE PRODUCTION AS A FUNCTION OF ANALYTIC PERCEPTION WITH A SPEECH PATTERN ELEMENT HEARING AID

○C TOFFIN, *K-E SPENS, □K SMITH, □R POWELL, ○P LENTE, □A FOURCIN, □A FAULKNER, *M DAHLQVIST, ○E FRESNEL-ELBAZ, #F CONINX, #C BEIJK, *E AGELFORS, □E ABBERTON
*KTH, SW □UCL, UK ○FONDATION ROTHSCHILD, FR #ST MICHELGESTEL, NL

ABSTRACT

The work described concerns an aspect of the way that real-time auditory feedback of a speaker's voice pitch information can have an appreciable influence on his or her control of vocal fold vibration in continuous speech. Larynx frequency range, vocal fold vibrational regularity and even detailed aspects of voice quality may, for a few profoundly deaf people, be markedly affected. The results discussed arise from the use of a phonetically motivated hearing aid, SiVo (Sine Voice), which has been used in controlled field trials in four countries with a total of 22 profoundly hearing impaired patients. The SiVo aid responds only to the voiced segments of speech and is designed to provide, for each separate input larynx period, a sine wave output which is matched to its user's residual hearing ability. In this first phase of work, the basic noise resisting neural network processing has been trained only on the use of targets produced by English speakers.

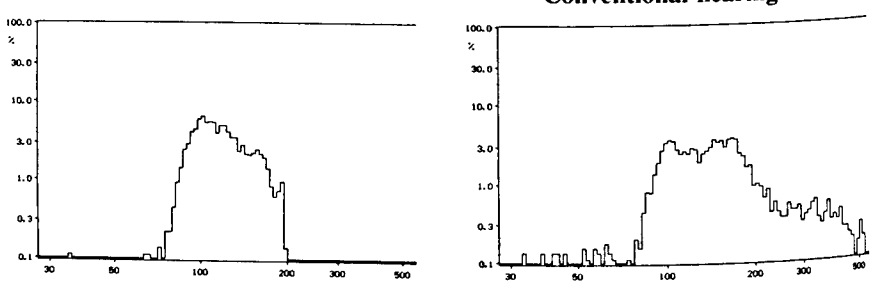
The training of the hearing impaired patients has been equally based on the balanced use of their SiVo and conventional aids and has involved their perception of intervocalic consonantal contrasts, single segment question / statement intonations, and SVO stress placement. No interactive speech production training based on the use of visually presented pattern element displays has been used for this phase of the work.

BACKGROUND

A suite of analytic programs has been applied to the quantitative assessment of the speech and laryngographic recordings obtained at the start and the end of trials in the different centres when the patients were using: their conventional hearing aids, the SiVo aid, and no aid at all. Special reference is made here for three patients, to the influence of these different conditions on: voice quality and vibratory regularity, and on the control of intensity and overall timing. In parallel with the speech based measurements, psycho-acoustic tests to assess temporal discrimination and frequency acuity were made in addition to standard pure tone based audiometry; these and speech receptive assessment results for the whole project are discussed in greater depth separately (Faulkner et al, these Proceedings).

For some of the production measurements major improvements associated with the use of the SiVo aid were found which were reversible even in a single recording session simply by changing back to a conventional hearing aid. An especially striking example of the influence of SiVo auditory monitoring on the speaker's larynx frequency range, Fx, is given below.

SiVo aid **Conventional hearing aid**



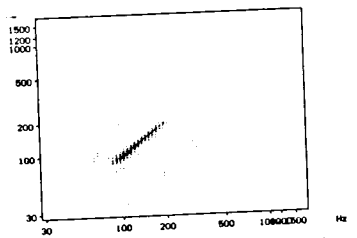
Larynx Frequency Distributions for MR

SiVo

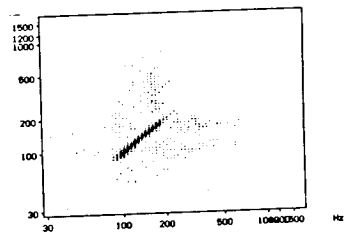
Larynx period detection and sine wave presentation at the most comfortable level for the user at all larynx frequencies, using in the ear audiometry and memory storage are provided within the SiVo aid itself. The recordings were made in the quiet and, in consequence, there was no benefit from the noise resistant features of this analysis.

C H A

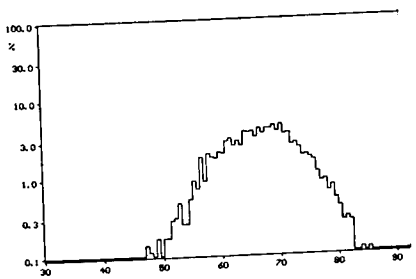
The Conventional Hearing Aid used by each patient in the study provided the best output possible for the individual user. In all cases in the work, the conventional hearing aid was far more familiar in use and sound quality, and equal amounts of training effort were devoted to both aids. All the results are a function of the aid used in the session.



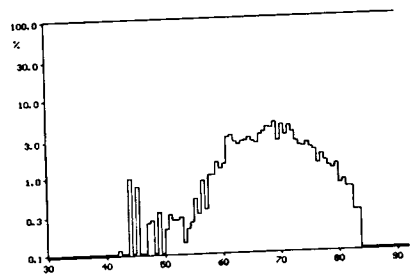
Larynx frequency crossplot using SiVo



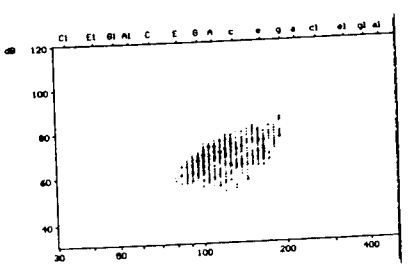
Larynx frequency crossplot with CHA



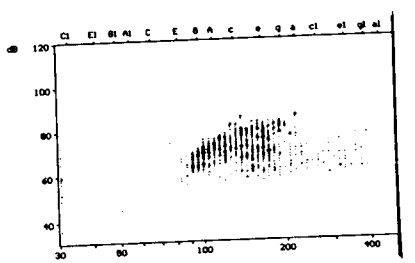
Speech Amplitude Distributions using:--



SiVo and CHA self monitoring



Sivo based second order Phonetogram



CHA based second order Phonetogram

PHYSICAL ANALYSES

The larynx frequency analyses shown on the first page, have been derived from period by period measurements. They are based on the synchronous speech and electro-laryngograph recordings which were made routinely during the clinical sessions. Both spontaneous and read passages were used, of at least two minutes duration. Only the read passages have been used here since this has made it possible for the speaker to aim for the same prosodic structures in the two monitoring conditions - SiVo and CHA - and for detailed segmental waveform comparisons to be made in subsequent analysis.

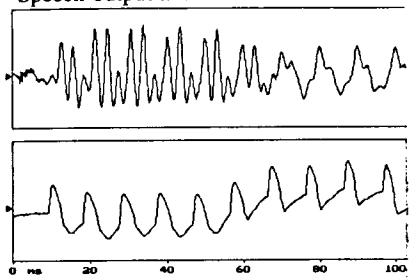
Speech excitation analyses are often based on windowed averages of more than one cycle of laryngeal, vocal fold, vibration. Larynx period, excitation epoch, based measurements for larynx frequency, Fx, give a more detailed foundation for the examination of speech waveforms. They also correspond to important aspects of many perceptually significant speech events. For example, the cycle-to-cycle period irregularities and breathy consonantal onsets of normal speech require this level of description for their detection and understanding; and in speech pathology temporal detail is useful in both assessment and training. In audiology, the ability of the hearing mechanism to base the sensation of pitch on the timing of acoustic events gives normal listeners the ability to hear creaky voice contrasts. For the very profoundly hearing impaired, the lack of peripheral frequency selectivity leaves essentially only temporal processing as the basis for their perception of voice pitch.

This is the case for the French patient MR, who lost his hearing (encephalitis after the essential stages of speech and language acquisition at 5y: 125Hz,65dB; 250Hz,80dB; 500Hz,105dB). For him, the conventional hearing aid can often give ill defined voice pitch periodicity information as a result of the varying harmonic structure of the acoustic input. SiVo analysis is designed to overcome this difficulty by the use of vocal fold closure detection training in the definition of its internal analysis algorithm and the provision of a sine wave acoustic output which, since it has only one harmonic component, does not change the nature of peripheral auditory temporal resonse with changes in input voice pitch.

The rather gross differences between the Fx distributions on the first page is an evident result of the lack of precise monitoring control afforded to MR by the conventional aid. The further analyses of his spoken outputs in the two monitoring conditions (for exactly the same recordings) give the basis for a more detailed understanding of what is happening. The larynx frequency crossplots simply show the distribution of successive pairs of Fx values, derived from the detection of successive epochs of excitation, throughout the whole speech sample. It is evident that it is not only the range of vocal fold frequencies which has been disturbed by the use of the CHA for monitoring but more importantly the temporal organisation at the quite detailed level of period to period closure. The speech amplitude distributions (probability plotted against dB) are based on the vocal fold synchronous determination of peak amplitude in each Fx period. They are remarkable for their similarity - apart from the low amplitude differences due to CHA induced creak. Overall loudness monitoring has not been changed by the switch from one aid to another. This is to be expected if it is essentially pitch perception which is not adequately supported by the conventional hearing aid.

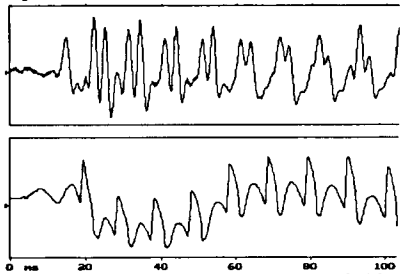
The final figures on the preceding page link the physical correlates of pitch and loudness through a development of ordinary phonetogram analysis. Here, once more, vocal fold closure detection has made it possible to obtain linked "instantaneous" measurements of Fx and excitation amplitude. In addition, the phonetogram analysis has only taken note of those pairs of successive vocal fold vibrations which have fallen into the same, quarter tone, analysis bin. In this way, the stable core of phonatory activity is shown and minor irregularities are eliminated. The SiVo based phonetogram is normal in shape for both frequency and amplitude. The CHA phonetogram is disorganised in the joint amplitude - frequency occurrences.

Speech output and Lx control with SiVo



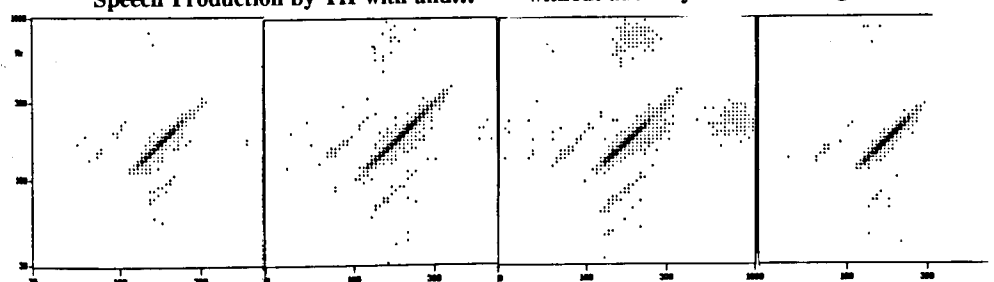
The use of the same read passage for all monitoring conditions has made quite detailed waveform level comparisons possible. The pair of waveforms above show the plots for speech, Sp above, and the laryngograph signal, Lx. The use of the SiVo aid to monitor his speech activity has not only enabled MR to control the broad levels of his phonation but also to produce essentially 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 Lx baselines).

Speech output and Lx control with CHA



MR has made many recordings of this passage during the course of the work and 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 is not able to provide him with an adequate sensation of voice pitch and in order parially to overcome this lack, he has a tendency to produce a breathy excitation spectrum - which has proportionately more energy at the low end of the spectrum. The Sp & Lx waveforms show this fairly clearly.

Speech Production by TH with and... without auditory feedback using SiVo



A complementary aspect of self monitoring speech with different methods of acoustic feedback comes from the comparison of the effect of the use of the 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, going from left to right: first used the aid to assist in reading a standard passage; then immediately after read again with no aid; then after an hour read again with no aid; and finally used the SiVo aid again for monitoring. At the beginning and end, using the SiVo for monitoring, the means of the Fx distributions are identical as are also the irregularity measures which indicate the departure of closure periods from the diagonal. In the absence of auditory control, there is both more variability between productions and more variability within. Mean & modal Fx values are different and, more obviously above, irregularity measures differ. This is not the case necessarily with the use of a CHA since here the same errors of control are repeatable.

We are grateful to our other colleagues; & to Laryngograph Ltd for its analyses.