SIVO-II: A SPEECH ANALYSING HEARING AID FOR PROFOUNDLY HEARING IMPAIRED PEOPLE

A Faulkner[1], J R Walliker[1], F Coninx[2], M Dahlquist[3], C Beijk[2], E Fresnel-Elbaz[4], K J Smith[1], J Wei[1], A Bosman[5], G F Smoorenburg[5], and A J Fourcin[1] for the STRIDE Consortium

Department of Phonetics and Linguistics, University College London [1], Institute voor Doven at St Michielsgestel[2], Department of Speech Communication and Music Acoustics, KTH[3], Fondation Rothschild (Paris)[4], University Hospital Utrecht [5]

ABSTRACT

The SiVo-II speech analysing hearing aid has been developed and assessed as part of the STRIDE project [1]. The project aimed to assess the potential of a speech pattern analysing mode of processing in a hearing aid for the profoundly hearing impaired. Trials in this user group in four European countries indicated that the speech analytic approach has significant advantages, especially for speech perception in noise

INTRODUCTION

Speech analysis has been proposed as an important component of signal processing in hearing aids for the profoundly impaired [2]. This approach has several potential advantages. It facilitates the matching selected speech information to residual hearing ability, and can make use of noise-resistant speech analysis. It could, moreover, lead to a commercial product that is highly cost-effective compared to alternatives such as cochlear implants

THE SIVO-II AID

The SiVo-II [3] is a wearable body-worn unit that allows speech information processing algorithms to be tested by profoundly hearing impaired subjects in daily life. It makes use of a TMS320C50 fixed-point DSP. The SiVo-II aid processes speech to extract voice frequency and amplitude information and match this to residual low frequency hearing. The signal presented by the aid is a sinusoid whose frequency and amplitude are controlled so as to preserve voice pitch and loudness information. Mapping of both frequency and intensity range are employed to ensure that the signal is always audible without discomfort.

To achieve optimal matching to residual hearing, the SiVo aid has been designed to act as its own audiometer. The audible intensity range at each frequency is determined using the user's standard transducer and ear-mould, ensuring correct calibration. This range is then directly used to control the intensity range of the aid's output.

Noise resistant fundamental frequency extraction

Fundamental frequency extraction is carried out by a Multi-Layer Perceptron (MLP) artificial neural network. The MLP algorithm has been trained to produce from the acoustic speech signal an output pulse corresponding to each instant of larynx closure, and hence, to give a cycle-by-cycle estimate of voice fundamental frequency. Training is achieved by adjusting the coefficients of the MLP to maximise the correspondence of its output to a target signal derived from an electrolaryngograph. It has proved possible to train the MLP to operate effectively with speech to noise ratios in the range of 5 to 10 dB, which are known to cause severe difficulty to profoundly hearing impaired users of conventional hearing aids.

In noise, the MLP method has been found to perform as well as other methods that are capable of implementation in a wearable processor [4,5]. For the classification of speech as voiced or voiceless, the MLP method was superior to all others. Since voicing information is a primary source of lipreading support, the MLP method was judged appropriate.

POLY-LANGUAGE ASSESSMENT TOOLS

One objective of the STRIDE project has been to provide assessment methods that are comparable between different languages and test centres. These included the following two tests that were employed in the user trials:

Vowel-Consonant-Vowel Tests

Vowel-consonant-vowel (vCv) tests of audio-visual consonant perception have been defined as a segmental basis for quantitative comparison across languages. Tests have been prepared using 10 consonants (**p b m f v t d s n l**) that are common to Dutch, English, French and Swedish.

Prosodic Test

In these same four languages, perceptual stress/accent can be cued by a major change in fundamental frequency on the contextually important word, and this provides the basis for a common approach to prosodic assessment. Tests have been designed using lists of three mono-syllabic word utterances with sentence stress on one of the three words cued by a falling pitch pattern..

USER TRIALS

Four clinical centres took part, each in a different country and using a different language, with a total population over the two phases of 34 participating hearing aid users. The current report concentrates on the phase II trials, in which 22 adult users took part. The selection criterion was a limited ability to make use of conventional hearing aids to aid lipreading in the quiet, and a profound post-lingual sensori-neural hearing loss. The age, duration of deafness and hearing losses of the selected group are shown in table 1.

Subjects received wearable SiVo-II aids and had conventional aids refitted where clinicians judged that the existing aid was not optimal. Speech perceptual assessments were carried out before and after a training period. Training was provided over a period of typically 6 to 8 weeks, during which users attended for three or four training sessions of approximately 2 hours duration. During each training session, matched training was given in the use of both the SiVo-II and the conventional hearing aid (CHA).

Table 1. Summary of hearing-impaired user group characteristics.

Country		FR	SW	NL	UK	All
	n	5	2	7	8	22
Age	*mean*	*59*	*56*	*53*	*48*	*53*
	min	*35*	*33*	*35*	*24*	*24*
	max.	*70*	*78*	*85*	*71*	*85*
Years of Deafness	*mean*	*42*	*33*	*23*	*23*	*28*
	min	*30*	*10*	*6*	*6*	*6*
	max.	*60*	*56*	*:50*	*53*	*60*
250 Hz HL (dB ISO)	*mean*	*87*	*83*	*78*	*86*	*83*
	min	*75*	*70*	*15*	*70*	*15*
	max.	*95*	*95*	*105*	*100*	*105*
0.5, 1 & 2 kHz HL (dB ISO)	*mean*	*106*	*99*	*104*	*110*	*106*
	min	*90*	*93*	*93*	*93*	*90*
	max.	*125*	*105*	*110*	*123*	*125*

Three tests of receptive ability were administered: vCv consonant lipreading with and without support from an aid; stress pattern recognition; and "connected discourse tracking" (CDT) using texts designed for language learners and of similar complexity across languages. CDT was performed during the training sessions, of which it formed an integral part. Specially designed questionnaires provided a more global evaluation.

The results show rather marked variation from one country to another, and over individual users within each country. While the results combined across the users in all four countries show the SiVo-II aid to be performing only at the same level as a conventional hearing aid, the UK subject group as a whole, and individual users in other countries, showed a clear advantage from the SiVo aid.

Objective measures of speech receptive benefit

Where data come from all four field trial centres, the dominant factor in each case is that of country. In the CDT and stress test, the aid effect depends significantly on the country.

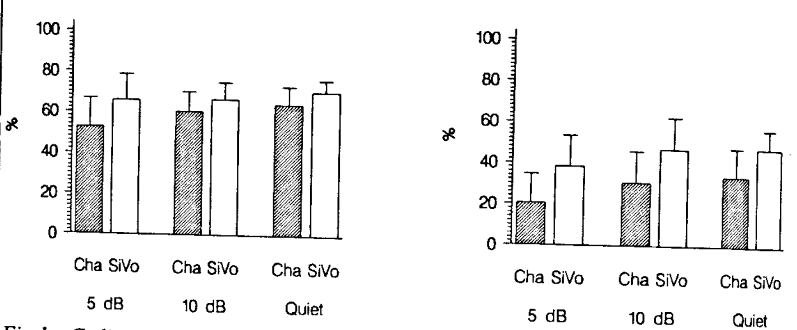

Fig 1. vCv lipreading in speech spectrum shaped noise. The figure shows the percentage correct score (left panel) and percentage of voicing information received (right panel) when using the SiVo-II aid (white bars) and the reference conventional hearing aid (hatched bars). Performance is shown for speech in quiet and speech-to-noise ratios of 5 and 10 dB. The error bars show the upper 95% confidence limit.

One unexpected factor in the results from different countries is now known to have been the sound reproducing loudspeakers used in Sweden and the Netherlands. The SiVo-II analysis algorithms are designed to operate with live speech, and require the preservation of the temporal characteristics of the speech waveform for correct operation. After results were collected, it was found that the SiVo analysis cannot operate properly with the acoustic signal from the loudspeakers used at IvD and KTH. As a result, the Swedish and Dutch data from tests using recorded speech (vCv and stress tests) at IvD and KTH cannot be taken to be representative of the benefit that would be expected using live speech.

vCv in noise (Fig 1)

The noise-resistance of the SiVo processing was expected to be a major factor in objectively measured benefit. vCv tests in noise were made in the UK. Data were collected at different background noise levels, and from a group of 11 SiVo-II users, three of whom were not in the STRIDE field trial group, but met the criteria for inclusion. There was a significant advantage from the SiVo-II aid compared to the CHA [$F(1,10) = 9.84$, $p = 0.011$]. There was also a significant effect of noise level [$F(2,17) = 4.63$, $p=0.025$]. At the poorer signal-to-noise ratio (5 dB) the scores obtained with the CHA were equivalent to scores obtained by the same subjects in unaided lipreading, that is, there was no benefit from the CHA. At this signal-to-noise ratio, however, the SiVo-II aid is still providing a useful degree of lipreading support. This advantage comes from the preservation of voicing information in the output from the SiVo aid at levels of noise that prevent the perception of this information from simple amplification of the noise+speech signal.

VCV in quiet

In quiet, there was no significant different between the SiVo-II and CHA across the group as a whole or in any of the four countries. Both aids significantly assisted lipreading. There was considerable individual variation, with individual users in each country showing greater benefit from the SiVo-II aid. The data from the Swedish and Dutch subjects are likely to have been affected by the interaction between the SiVo speech analysis and loudspeaker performance.

CDT

The UK group scored significantly higher in CDT with the SiVo-II than the CHA [$F(1,7) = 5.64$, $p<0.05$]. The other user groups showed no significant difference between the two aids.

Stress placement

Once more the UK group scored significantly higher with the SiVo-II than the CHA [$F(1,7) = 8.33$, $p<0.025$]. The other user groups showed no significant difference between the two aids. The main sources of variation in the overall data were country [$F(3,17) = 20.82$, $p<0.005$] and an aid*country interaction [$F(3,17) = 8.3$, $p<0.01$]. There were also effects due to voicing [$F(1,16)=6.3$, $p<0.025$], where phrases using voiceless consonants gave higher scores higher than all voiced phrases, and an aid*voicing interaction [$F(1,16) = 6.6$, $p<0.025$].

Subjective Evaluation

After training, users completed a questionnaire. Overall, users were slightly more satisfied with their conventional hearing aid (CHA) than with the SiVo. This outcome is largely due to the Dutch users, who on the whole preferred the CHA. The Swedish and French users preferred the SiVo-II, and UK subjects showed no clear preference. Considering the difference in convenience of use between a post-aural conventional aid and the bulky SiVo-II, this is an encouraging outcome.

Users were asked whether they would wish to use a SiVo-based hearing aid, either in its current form, in a smaller package that also included an amplification mode. Fifteen of the 22 users taking part in the trial expressed the wish to use a smaller aid that offered both types of processing.

CONCLUSIONS AND FURTHER DEVELOPMENTS

The STRIDE user trial has shown that the approach has clear benefits for some users. The speech analytic approach appears to be especially useful for providing lipreading support in noise. From the data gathered in the project, it is estimated that up to 160,000 persons may be able to benefit from an aid that incorporates both speech pattern extraction and amplification.

Improvements in the performance and design of speech pattern extracting aids are required before an acceptable product could be developed. The sensitivity of the SiVo analysis algorithms to loudspeaker performance also requires to reduced, so that television and other domestic sound reproduction equipment can be used more effectively by users of such aids.

The new EC TIDE project OSCAR is now in progress, and will address these issues. A new version of the aid, SiVo-III, has been developed, which also encodes voiceless excitation information. This has been shown to significantly add to the lipreading support available from fundamental frequency and amplitude information [6]. The OSCAR project is also examining the potential of tactile and combined auditory/tactile aids using speech analytic processing.

ACKNOWLEDGEMENT

Supported by EC TIDE project TP133/206 (STRIDE) and UK MRC grant G9020214. Other partners in the STRIDE project were: Laryngograph Ltd, Oticon A/S, and CNRS Parole et Langage at Aix-en-Provence, CCA Wagram (Paris).

REFERENCES

[1] STRIDE (1995) "Final report of TIDE project TP133/206 STRIDE", ref.: STRIDE/1995/2, Dept Phonetics and Linguistics, University College London

[2] Fourcin, A. J. (1990). "Prospects for speech pattern element aids," Acta Otolaryngol. (Stockh), Suppl. 469. 257-267.

[3] Walliker, J., Daley, J., Smith, K., Faulkner, A. and Fourcin, A. (1993) "Speech analytic hearing aids for the profoundly deaf: Technical design aspects and user field trials results" In: B. Granstrom, S. Hunnicutt and K-E. Spens (Eds) *Speech and language technology for disabled persons,* ESCA/KTH, Stockholm, pp 35-43

[4] Bosman AJ and Smoorenburg, GF (1994) "TIDE project 206 Internal Report. "Evaluation of Three Pitch Tracking Algorithms", ref.: STRIDE/1994/2. Dept Experimental Audiology, University Hospital Utrecht,

[5] Wei J, Howells D, Fourcin AJ, and Faulkner A. (1993) "Larynx period and frication detection methods in speech pattern hearing aids" Proc. Eurospeech '93, 3rd Eur. Conf. on Sp. Comm. and Tech. Vol. 3. ESCA, pp 2037-2049.

[6] Rosen, S., Faulkner, A., Reeve, K. and Smith, K. (1995) Voicing, fundamental frequency, amplitude envelope and voicelessness as cues to consonant identity. Proc. ICPhS, Stockholm, 1995