European Conference on
Speech Technology
Edinburgh, Scotland, UK
September 1987

ARTIFICIAL CORRECTIONS TO DEAF SPEECH AND THE DEVELOPMENT OF VISUAL SPEECH TRAINING AIDS (1)

B. Maassen[2], N. Arends[3], D.Povel[3]

ABSTRACT

The causes of the (low) intelligibility of deaf speech were studied, following a so called "Speech Transformation Method". By manipulation of analysis parameters, errors of articulation, intonation and timing, that occurred in sentences spoken by deaf children, were artificially corrected one-by-one. Intelligibility tests showed that suprasegmental corrections caused only a small improvement (from 24% to 34% word-intelligibility), whereas segmental correction increased intelligibility scores to 74%. The results of this study formed one of the starting points for a research project aiming at the development of a visual speech training aid: a device that supplies visual information on the acoustic quality of deaf speech.

1. INTRODUCTION

In general, the speech of the deaf is difficult to understand for naive listeners. In section 2 of this paper, a series of diagnostic studies is presented, in which the contribution of various deviating characteristics in the speech of deaf people to their intelligibility was determined. In section 3 a training device, a so called Speech Visualizer (SV) is described: a minicomputer that transforms acoustic speech input into graphics output.

2. ARTIFICIAL CORRECTIONS TO DEAF SPEECH

Aim of the study

Several investigators [1] have shown that deaf speech deviates considerably from speech of normally hearing persons in all three traditionally distinguished acoustic aspects: temporal structure, intonation contour, and segmental quality. Aim of the present study was to determine the relative contribution of each of these errors to overall speech intelligibility. For this, use was made of a so called "Speech Transformation Method", by which deviating characteristics were corrected one-by-one. The corrections were evaluated by means of intelligibility tests using naive listeners.

Corrections of Deaf Speech

Thirty sentences spoken by ten prelingually deaf children, aged 12 to 14 years, were recorded. The same sentences were also spoken by two normally-hearing children, one boy and one girl, which served as model sentences for the corrections to be

---

[1] The research reported was supported by a grant from the Netherlands Organization for the Advancement of Pure Research (ZWO) and the Institute for the Deaf, St.Michielsgestel.

[2] Inst.Medical Psych., Univ.Hospital Nijmegen, PB 9101, 6500 HB Nijmegen.

[3] Dep.of Exp.Psych.,Post Box 9104,6500 HE Nijmegen, The Netherlands.

10.21437/ECST.1987-143

carried out. All 60 sentences were LPC analyzed and temporally measured. Deaf sentences were corrected by transforming analysis parameters in the direction of the model sentences with respect to each of the following acoustic dimensions, both separately and combined: temporal structure, intonation (fundamental frequency) and segmental quality (spectral envelope as described by reflection coefficients). This yielded the corrected versions (1) to (7), listed in Table 1.

Testing Intelligibility

Together with the unmanipulated, resynthesized deaf sentences, version (0), the seven versions of each sentence were distributed over stimulus tapes according to a randomization design based on Latin Squares (footnote 4). The tapes were played to different groups of 16 to 20 listeners with no earlier experience with deaf speech. Subjects listened to each sentence twice, after which they wrote down as much as they understood. Intelligibility scores were expressed as percentages of words correctly understood.

Results

Table 1 shows the mean and median intelligibility scores for all eight versions.

Table 1. Mean intelligibility scores and standard errors (s.e.) of the original (0) and seven corrected (1) - (7) versions of 30 deaf sentences.

		M	s.e.
(0)	resynthesization	20.7	4.47
(1)	temporal correction	25.8	4.72
(2)	intonation	27.2	5.78
(3)	vowels corrected	47.3	6.36
(4)	stops,fric.,affr.	28.5	5.53
(5)	all sounds corrected	71.4	5.96
(6)	intonation+temp.str.	25.2	5.54
(7)	complete corr.(5)+(6)	85.8	4.33

Both types of suprasegmental corrections: temporal structure and intonation, yielded only small improvements in intelligibility, which, however, were significant in an analysis of variance ($p<0.05$). Segmental correction, on the other hand, caused a dramatic, highly significant ($p<0.001$) improvement; for the major part this increase was obtained by correcting vowels ($p<0.01$).

Discussion

The data clearly suggest that, in order to further improve deaf speech intelligibility, emphasis should be on training segmental aspects; especially correct articulation of vowels deserves more attention. Second, only when segmental aspects are sufficiently developed, efforts toward improving the suprasegmental features are justified, the interaction being significant ($p<0.01$).

One qualification should be added. Transposing reflection coefficients from model to deaf sentences to correct speech sounds also improves voice quality. This may have led to an

[4] For more detailed information on these versions, the procedure of the intelligibility experiments and statistical analyses, the reader is referred to [3, 4, 5, 6].

overestimation of the segmental corrections, or, reversing the argument, voice quality might be an important determiner of (deaf) speech intelligibility.

3. SPEECH TRAINING AIDS FOR THE DEAF

Aim of the project

Recently, a research project has been initiated that aims at the development of a visual speech training aid, the so called "Speech Visualizer" (SV). The SV will be introduced in the speech trainig curriculum of deaf-born children from as early an age as possible. Based upon the study presented in the previous section, the SV will supply information on those aspects of speech that are often realized incorrectly by deaf speakers, and, moreover, are detrimental to intelligibility. Thus, emphasis will be on computer programs giving feedback on voice quality and correct vowel articulation [9]. A thorough discussion of the remaining starting points of the project can be found in [7, 8]. Let us mention just three of them, that have direct implications for the hardware and analysis techniques of the SV:

1. Emphasis will be on the <u>product</u>, not the process of speech, for the following reason. Deaf language learners, typically lacking auditory information, do have intact visual, tactile and kinesthetic senses, which supply feedback on articulation, the process of speaking. Thus, the SV will be mainly based on analyses of the acoustic speech signal. We will, however, use an Electro-Glottograph (ELG), a non-invasive device to measure voice-quality and pitch.

2. Design of the SV relates to the decomposition model of speaking [2]. That is, separate programs will be developed to give feedback on the basic speech skills: respiration, phonation and articulation.

3. All relevant aspects of speech will be simultaneously calculated, even if they do not make part of the display at a particular moment. Background monitoring of non-displayed aspects in combination with a hierarchical organization of functions, makes the system robust as a stand-alone training device, thereby increasing opportunity for drill-practice.

Computer Configuration

The prototype of the SV, as it currently operates, consists of a host computer based upon a Motorola 68000 microprocessor, interfaced to a custom built pre-processing board (see Figure 1). The latter is designed around a Texas Instruments TMS32010 chip, which can perform real-time analyses of incoming speech signals. The pre-processing board further contains a Hitachi HD63701 processor, program memories, a 16 kWord data memory, two analog-digital (A/D) converters, such that two signals (acoustic speech signal and ELG signal) can be processed simultaneously, and a digital-analog (D/A) converter. The host computer contains 1 Mbyte of internal memory, two external memories: a floppy disk and a hard disk, a graphic interface and a terminal. The host computer runs under the OS-9 operating system.

Displays of intonation, loudness and voice quality will be shown at the conference.

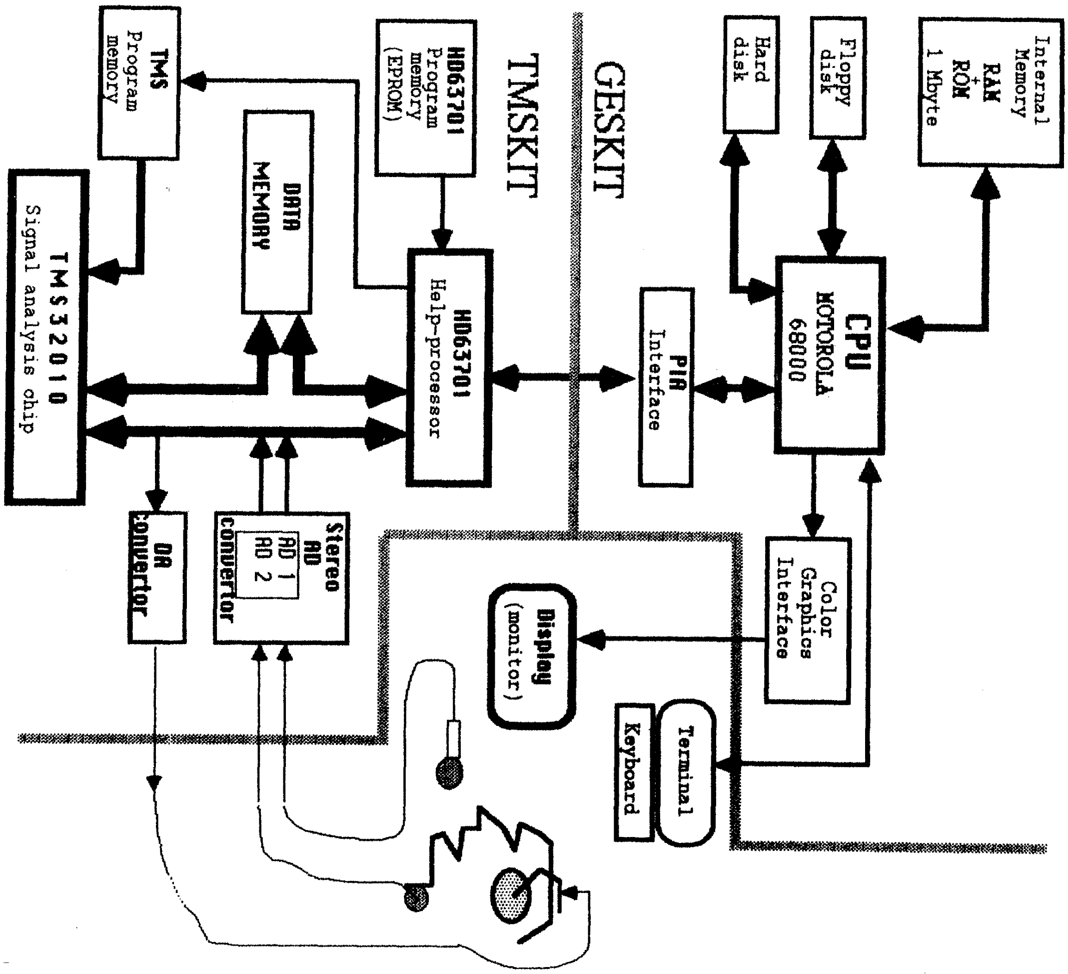

Figure 1. Block diagram of the Speech Visualizer.

References

[1] Hudgins, C. V., & Numbers, F. C. (1942). An investigation of the intelligibility of the speech of the deaf. Gen.Psych.Monogr., 25, 289 - 392.

[2] Ling, D. (1976). Speech and the Hearing Impaired Child. Alex.Graham Bell Ass. for the Deaf, Inc.

[3] Maassen, B. (1985). Artificial Corrections to Deaf Speech. PhD dissertation, University of Nijmegen.

[4] Maassen, B. & Povel, D.J. (1984a). The effect of correcting temporal structure on the intelligibility of deaf speech. Speech Communication, 3, 123 - 135.

[5] Maassen, B. & Povel, D.J. (1984b). The effect of correcting fundamental frequency on the intelligibility of deaf speech and its interaction with temporal aspects. J.Ac.Soc.Am., 76 (6), 1673 - 1681.

[6] Maassen, B. & Povel, D.J. (1985). The effect of segmental and suprasegmental corrections on the intelligibility of deaf speech. J.Ac.Soc.Am., 78, 877 - 886.

[7] Povel, D.J. (1987). Enkele theoretische overwegingen bij de ontwikkeling van visuele hulpmiddelen ten behoeve van het spreekonderricht aan dove kinderen. Tijdschr.Logop.Audiol., 17-2.

[8] Povel, D.J. and Maassen, B. (1987) Visual information and speech acquisition of the deaf. Eleventh Intern.Congr.Phon.Sciences, Tallin, USSR, August 1-7, 1987.

[9] Povel, D.J., & Wansink, M. (1985). A computer-controlled vowel corrector for the hearing-impaired. J.Sp.Hear.Res.,29, 99-105.