1

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS
STANDARD REFERENCE MATERIAL 1010a
(ANSI and ISO TEST CHART No. 2)

National Library
of Canada

Bibliothèque nationale
du Canada

Canadian Theses Service

Service des thèses canadiennes

Ottawa, Canada
K1A 0N4

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

**A Digital Signal Processing-Based Hearing
Prosthesis and Implementation of Principal
Components Analysis for a Tactile Aid**

by

Timothy J. Rahrer, B.Eng.

A thesis submitted to the
Faculty of Graduate Studies and Research
in partial fulfillment of the requirements
for the degree of
Master of Engineering

Faculty of Engineering
Department of Electronics
Carleton University

October, 1990

© copyright
1990, Timothy J. Rahrer

National Library
of Canada

Bibliothèque nationale
du Canada

Canadian Theses Service Service des thèses canadiennes

Ottawa, Canada
K1A 0N4

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-315-68898-X

Canada

Abstract

In an effort to provide improved assistive devices to the auditorily impaired, some research prosthetic systems use complex speech processing algorithms and often allow customization to the individual user. This thesis documents the development of a programmable, modular speech processor hardware based on a commercial digital signal processing microprocessor. In addition, an input / output hardware module was developed for a vibrotactile output application. A detailed literature review indicated that the principal components analysis (PCA) of speech spectra may provide information suited to the needs of a tactile aid wearer. Therefore, the PCA algorithm was selected for the tactile aid and implemented in real-time on the speech processor hardware.

Acknowledgements

I would like to thank my supervisors, Tad A. Kwasniewski, and Brian W. Tansley for their patience, encouragement and support, both moral and material. In addition I wish to acknowledge the Department of Electronics, including staff and fellow graduate students for providing an enjoyable environment in which to work. I would also like to thank Vanda for her patience and understanding during the course of this project.

Finally, I would like to dedicate this work to my parents.

Table of Contents

Abstract	iii
Acknowledgements	iv
Table of Contents	v
List of Figures	vii
List of Tables	ix
Glossary	x
 Chapter 1	 1
1.0 Introduction	1
1.1 Intent and Outline of the Thesis	5
1.2 The Problem	6
1.3 Possible Solutions - Introduction	9
1.3.1 Visual Displays	10
1.3.2 Cochlear Implants	12
1.4 Design Requirements of a Tactile Aid	15
1.4.1 Speech Processing Considerations - Speech Production	16
1.4.2 Speech Processing Considerations - Speechreading	20
1.4.3 Other Design Issues	24
 Chapter 2	 26
2.0 Introduction	26
2.1 Survey of Tactile Aid Research	26
2.1.1 Single Speech Feature Tactile Aids	28
2.1.2 Multifeature Tactile Aids	35
2.1.3 Summary of Tactile Aid Research	51
2.2 Characteristics of the Tactile Sensory Channel ...	57
2.2.1 Electrical and Mechanical Stimulation	58
2.2.2 Single Source Tactile Channel Characteristics	62
2.2.3 Multi-Source Tactile Channel Characteristics	69
2.2.4 Tactile Channel Capacity Conclusions	72
2.3 Speech Processing Strategy	73
 Chapter 3	 77
3.0 Introduction to Principal Components Analysis	77
3.1 Principal Components Analysis (PCA)	77
3.2 Principal Components Analysis of Speech	82
3.3 Configuration of PCA for a Tactile Aid	91
3.3.1 Spectral Analysis Variables	92
3.3.2 Principal Component Analysis Variables	94
3.5 Summary of PCA for Use in a Tactile Aid	96
 Chapter 4	 97
4.1 System Level Outline	97
4.2 Hardware	99
4.2.1 Digital Hardware	101

4.2.2	Analog I/O Hardware	112
4.3	System Software	122
4.3.1	Spectral Analysis	123
4.3.2	Spectral Energy Detection and Conditioning	144
4.3.3	Principal Components Calculation	152
4.3.4	Software Module Integration and Statistics	152
4.4	System Integration and Testing	154
4.4.1	Tactile Psychophysical Testing	154
4.4.2	Hardware / Software Integration	159
4.4.3	System Function Verification	161
4.4.4	Final Tactile Aid Considerations	163
Chapter 5		165
5.0	Conclusions	165
5.1	Future Work	170
5.2	Concluding Remarks	174
Appendix I		175
Appendix II-A		176
Appendix II-B		186
Appendix II-C		190
Appendix II-D		191
Appendix II-E		196
Appendix II-F		205
Appendix III		208
Appendix IV-A		210
Appendix IV-B		219
Appendix IV-C		221
Appendix IV-D		222
Appendix V		231
References		237
Bibliography		253

List of Figures

Figure 1 Human Communication System - Transmitter	2
Figure 2 Human Communication System - Receiver	3
Figure 3 Tactile Vocoder Configuration	37
Figure 4 Complex visual pattern (after Scott, 1980)	46
Figure 5 Transformation of Fig. 4, to organization by size	47
Figure 6 SRA-10 Tactile Speech Reception Aid (Scott, 1980)	49
Figure 7 Equal Sensation Curves (Verillo, 1970)	65
Figure 8 Three Dimensional Rotation Example	79
Figure 9 Principal Component Basis Vector #0	83
Figure 10 Principal Component Basis Vector #1	84
Figure 11 Principal Component Basis Vector #2	84
Figure 12 Principal Component Basis Vector #3	85
Figure 13 Principal Component Basis Vector #4	85
Figure 14 Variance Accounted For By Principal Components	87
Figure 15 Functional Module Diagram	99
Figure 16 TMS320c25 Architecture (TMS320c25 User's Guide, 1989)	104
Figure 17 Signal Processing Hardware Blocks	107
Figure 18 Program and Data Memory Maps	108
Figure 19 Address Decoding Circuit	109
Figure 20 READY Signal Generation Circuit	110
Figure 21 TLC3204x Interface	116
Figure 22 Transducer Evaluation Apparatus	118
Figure 23 Transducer Frequency Response	119
Figure 24 Output Interface Circuit	121
Figure 25 Software Functional Block Diagram	123
Figure 26 FIR vs IIR Magnitude Response	127
Figure 27 DFDP CH9 Magnitude Plot and ANSI 1966 S1.11	130
Figure 28 Direct Form IIR Signal Flow Diagram	132
Figure 29 Cascade Structure	135
Figure 30 Parallel Structure	135
Figure 31 Direct Form II Second-order Filter	138
Figure 32 Transpose Form of a Second-Order Filter	140
Figure 33 TMS320c25 Program for an IIR Bandpass Filter	142
Figure 34 1 kHz Octave Frequency Response	143
Figure 35 Analog Energy Detector Scheme	145
Figure 36 Channel 9 Filter/Detector Response to Sudden Sinusoid Input	147
Figure 37 $20(\log(x))$ and 12 Segment Linear Approximation	149
Figure 38 Log Function and Linear Approximation Expanded	150
Figure 39 TMS320c25 Log Approximation and Error vs the Floating-point Log Function	151
Figure 40 Software Task Distribution	153
Figure 41 Tactile Amplitude Perception Test Apparatus	156

Figure 42	Tactile Amplitude Threshold Results	157
Figure 43	Maximum Comfortable Level Results	157
Figure 44	PC0 for Utterance 'adda'	162
Figure 45	PC1 for the Utterance 'sha'	163
Figure 46	Program Memory Decode Circuit and Timing	177
Figure 47	Data Memory Decode Circuit and Read Timing	180
Figure 48	Data Memory Write Timing	183
Figure 49	READY Signal Generation Circuit	185
Figure 50	Speech Processor Board Schematics - CPU	187
Figure 51	Speech Processor PCB Layout	190
Figure 52	Logic Analyzer Display of Write Followed by Read	192
Figure 53	Data Memory Address Decode Propagation Delay	194
Figure 54	Program Memory Access Timing form Logic Analyzer	195
Figure 55	Program Memory Address Decode Propagation Delay	195
Figure 56	Analog Driver Circuit	206
Figure 57	Analog Driver Circuit	207
Figure 58	Frequency Response of Transducer #1	208
Figure 59	Frequency Response of Transducer #2	208
Figure 60	Frequency Response of Transducer #3	209
Figure 61	Frequency Response of Transducer #4	209
Figure 62	Bandpass Filter Bank Frequency Response	220
Figure 63	Linear Approximation of a Log Curve Segment	222

List of Tables

Table 1 Rates of Profoundly/Severely Deaf in North America	7
Table 2 Visemes and Corresponding Phonemes	23
Table 3a Summary of Tactile Aids Reviewed	53
Table 3b Summary of Tactile Aids Reviewed (Continued)	54
Table 4 Sensitivity of Body Sites (Kenshalo, 1978)	64
Table 5 Filter Bank Centre Frequencies and Bandwidths	129
Table 6 Filter Bank Characteristics	144
Table 7 Program Statistics	154

Glossary of Terms

- /*/** - Notation used in speech sciences to represent various speech sounds; ex. /b/ as in bat
- AM** - Amplitude modulation
- auditory nervous system** - Complex electrical pathway linking the sound-to-electric transducer (cochlea) of the ear to the brain
- carrier signal** - Original waveform, to which modulation has been applied
- CMOS** - Complementary Metal-Oxide Semiconductor. A manufacturing process and technology used to produce a class integrated circuits having low power consumption.
- cochlea** - Fluid filled, spiral chamber in the inner ear containing approx. 40,000 hair cells that convert mechanical sound energy to electrical signals
- CVC** - Notation indicating that the speech sample is of the form consonant - vowel - consonant.
- dB** - Decibel, unit of speech energy or power measurement.
- DFDP** - Digital Filter Design Package, software running on the IBM - PC, available from Atlanta Signal Processors, Inc.
- difference limen** - Minimum distance between two stimuli, so that they may be perceived as separate.
- DSP** - Digital signal processing or processor.
- electrotactile** - Stimulation of the sense of touch using electric current
- FM** - Frequency modulation
- formants** - Frequency peaks in speech sounds created by resonances in the vocal tract.
- fricative** - A class of consonant sound in which a constriction between two articulators causes a turbulence in the air flow through the vocal tract. Examples are /s, v, f, z/.
- fundamental frequency** - The frequency of vibration of the vocal chords, sometimes called pitch. Ranges in frequency from 100 - 300 Hz, and is higher for female speakers.

glue logic - Refers to the electronic hardware required to interface two or more VLSI integrated circuits together. Usually glue logic handles conversion of communication protocols and/or timing of signals between two ICs.

intonation - Changes of the fundamental frequency (pitch) of speech which often indicate emphasis

KLT - Karhunen-Loeve Transform or expansion, equivalent to principal components analysis.

Kynar - a Poly-Vinyl DiFlouride (PVDF) peizoelectric material available from the Pennwalt Corp., USA.

labial - Sound formed by lip opening / movement

manner of articulation - Method used in the vocal tract to produce the speech sound; ex. relaxed vocal cords produce 'voiceless' sounds

masking - Interaction between two or more stimuli, that changes the perception of one or both.

nasal - A class of consonants in which oral cavity is blocked and the air passes through the nasal cavity. Examples are /m, n/.

neurons - Specialized cells that transmit information throughout the body.

PCA - principal components analysis

PCB - Printed circuit board.

phoneme - The shortest speech sound that indicates a difference between two utterances (cf Appendix I).

place of articulation - Location(s) within the oral cavity where the airflow is controlled to produce a speech sound; ex. at the lips for /p/ as in put

plosive - A class of consonant sounds produced by a complete blockage of the oral cavity, followed by a sudden release of the built up pressure. Also called a stop consonant. Examples are /p, t, k, b, d, g/.

prelingual - Prior to the acquisition of speech production skills

prosody - Rhythm and stress patterns of speech or sounds

somatosensory - The tactile sense, the sense of touch.

spatiotemporal - Having to do with both time and space

switched capacitor - Refers to a signal processing technique in which capacitors are switched in and out of a circuit by a common clock, in order to produce the desired signal transfer function, ie filtering.

tactile - The sense of touch

tactile aid - Refers to a prosthetic device that converts sound to the sense of touch, allowing a deaf individual to understand speech through the sense of touch

tinnitus - 'Ringing' sensation in the ear without sound

tracking task - A common training and evaluation task in which the subject repeats or 'tracks' the what is spoken by the teacher. Scores for normally hearing subjects are about 95 - 110 wpm

vestibular - Refers to the system of fluid filled canals within the ear, that assists the brain in providing a sense of balance and equilibrium

vibrotactile - Stimulation of the sense of touch using vibration

viseme - Sounds that appear similar, visually when lipreading

VLSI - Very large scale integration, refers to modern electronic integrated circuits or 'chips'

vocoder - Voice coder, refers to a speech processing method in which the spectral energy of the sound is extracted and coded for transmission.

voicing - Modulation of the air flow from the lungs by the vocal cord vibrations.

wait-states - Refers to the number of clock cycles a microprocessor must wait while accessing a slow peripheral or memory.

Chapter 1

1.0 Introduction

The development of a prosthetic device for the handicapped can be a challenging task. This thesis discusses the design of an electronic speech processor and implementation of a tactile speech reception aid on this hardware, to assist those with the handicap of profound deafness. Figures 1 and 2 illustrate the problem at hand in a communication system context.

Figure 1 depicts the speech production mechanism, otherwise known as the message encoder / transmitter of the communication system. In Figure 1, the message to be transmitted is formulated in the brain and nerve signals are sent to control the vocal tract apparatus. Air expelled from the lungs forms a carrier for the speech information that can be either a complex repetitive wave formed by vocal cord vibrations, or a breath noise produced as the air is forced through a constriction in the vocal tract. The spoken message is created by the selection of carrier type, carrier frequency and by acoustic modulations of the carrier signal.

The modulations are created by muscular control over the vocal tract shape and openings. The resulting modulated sound travels through air (the transmission channel for this communication system) to the listener (receiver), as well as, to the speaker as feedback to control the transmission.

Figure 1 Human Communication System - Transmitter
(after Dudley, 1976)

Figure 2 illustrates the receiver of the human communication system, consisting of the peripheral signal conditioning (outer ear, ear drum etc.), the detector (cochlea) and message decoder (brain). Sound channeled into the ear causes the ear drum to vibrate. A mechanical linkage (via the ossicles) to the cochlea produces motion of a fluid inside the cochlear spiral chamber. This fluid motion causes hair cells within the cochlea to move, which in turn alters their electrical activity. Thus the cochlea may be thought of as a mechanical to electrical signal converter. Moreover the

cochlea is tonotopically mapped, meaning the production of electrical signals at different locations in the cochlea represent different frequencies of the incoming sound. A complex electrical pathway, the central auditory nervous system, carries the signals to the brain for message

Figure 2 Human Communication System - Receiver
(after Ainsworth, 1976)

decoding. Although not shown in Figure 2, the central auditory nervous system consists of an intricate network of pathways. The reader is referred to Stevens and Davis (1983), for specific details.

Profound deafness can be likened to damage of the receiver in a communication system. Because auditory feedback is needed for normal speech development, profound deafness also hinders speech production (i.e. "damages" the transmitter of the communication system). Most types of profound deafness involve damage to the electrical connection (auditory nerve) or to the cochlea (hair cells) directly. A prosthetic device may decode the modulated message-bearing signal (speech), extract the important information and recode this information in a form suitable for an alternate transmission channel to the message processor (brain). Herein lie the two greatest challenges:

- i) determining and extracting the information required to understand the message in the speech signal and
- ii) recoding this message in a form suitable for an alternate channel.

The success of any sensory prosthetic device is strongly dependant on the proper application and interface of technology to the requirements of the user. A plethora of elegant electronic circuitry and real-time software are useless if designed without a thorough understanding of the problem; in this case profound deafness. Therefore Chapters 1 and 2 discuss the problems associated with profound deafness, the resulting requirements for an assistive device and the important past research in the area of tactile aids for the profoundly deaf. With the resulting interdisciplinary knowledge in hand, the development of an electronic sensory aid may proceed. A glossary is provided to assist those

readers who are unfamiliar with either the electrical engineering or auditory prosthetics terminology.

1.1 Intent and Outline of the Thesis

The intent of this thesis is to develop a flexible, compact electronic testbed system for the evaluation of various real-time speech processing algorithms. This electronic system would replace the damaged portions of the receiver in the communication system depicted in Figure 2 and extract the message bearing features of the speech signal. A second intent of the thesis is the implementation of a tactile sensory aid using the testbed system, to demonstrate a sample application with recoding of the speech features to an alternate sensory transmission channel. To this end, the difficulties in speech reception and production, due to profound deafness and a brief review of auditory prosthetic methodologies, are discussed in the remainder of Chapter 1. The specific speech production deficiencies, as well as the limitations of lipreading outlined in the first chapter, provide the starting point in the design of a tactile aid hardware and algorithms. A review of important research in the area of tactile prostheses is undertaken in Chapter 2.

A review of specific aids in Section 2.1 contributes further insights into the design requirements. Section 2.2 discusses the information capacity of the tactile sensory channel and its limitations for use as a substitute for the auditory pathway. The design requirements outlined in Chapter

1, the review of past research discussed in Section 2.1 and the characteristics of the tactile sensory channel in Section 2.2, suggest that the principal components analysis (PCA) of speech spectra may be useful in a tactile aid, as proposed in Section 2.3.

Chapter 3 provides a description of PCA, reviews past research in the area, discusses the advantages of such a processing scheme and details the configuration of PCA for use in a tactile aid for the profoundly deaf. Chapter 4 describes the system development and implementation; Section 4.1 at a functional level, 4.2 - the hardware realization and 4.3 the corresponding software. Section 4.4 deals with the integration of the software and hardware into a stand-alone system. Scaling of the tactile output levels to suit the results of tactile vibratory testing is discussed. Sample system responses to speech input are also presented. Although subjective evaluations of the algorithm are beyond the scope of this thesis, proposals are made in Chapter 5 for this work. Chapter 5 also summarizes the work accomplished and provides suggestions for future research using the system developed.

1.2 The Problem

Although the term 'deaf', is used to refer to many degrees of hearing impairment, deafness usually refers to the inability to hear and understand speech. In cases of deafness where the standard corrective measures, i.e. amplification of

sounds, are not sufficient, the hearing impairment is designated as being profound. Hearing impairment is the single most prevalent chronic physical disability in North America (Smith-Kettlewell Institute, 1977) and as Table 1 illustrates, the number of profoundly deaf (loss > ~70 dB) individuals in North America, is significant.

Country	Rate/100,000 Age at onset <19	Rate/100,000 All Ages at onset	Source
Canada	na	> 800	a, 1989
USA	203	3237	b, 1974

a = Canadian Hearing Society (1989)

b = Schein and Delk (1974)

Table 1 Rates of Profoundly/Severely Deaf in North America

For the profoundly deaf person, the speech of others is perceived mainly through lip-reading, often termed speechreading. Speechreading involves interpreting the facial expressions, gestures and articulatory movements of the speaker in order to understand the speech. There are however, several factors which limit the effectiveness of speechreading. Many sounds, such as /p/ (in push) and /b/ (in bush) cannot be discriminated by lip-reading since they are produced with identical lip and jaw movements (Breeuwer and Plomp, 1984). Another problem exists in that many sounds do not produce clearly visible lip and jaw movements at all (Berger, 1972). Also the intonation and stress patterns

(prosody) of the speech signal, which are can be useful in speech perception, cannot be derived from speechreading. Prosodic features are important for deciding whether a sentence was spoken as a statement or a question, for establishing the urgency of the message, and for determining the word or words of emphasis in the sentence.

Profoundly deaf persons also lack access to the important auditory information about their environment. Warning and signalling noises such as car horns, sirens, timer buzzers, alarm clocks, doorbells etc., cannot usually be perceived. The inability to hear some signalling cues, such as an alarm clock, are inconveniences which can be overcome by devices which turn a light on and off, i.e. translate the auditory signal into a visual one. Not being able to hear a car horn however, may put the deaf person's safety at risk.

The problems of profound deafness become particularly acute for children, especially those who are prelingually (before learning to speak) deaf. Since the production of speech is closely related to the reception of speech, prelingually deaf children have an exceedingly difficult time learning to communicate orally. Most four-year-old hearing children have a complete working knowledge of their language, which is absorbed and processed without formal training. In comparison, most deaf children of the same age have only a limited grasp of oral communication, which has been attained through prolonged and intensive training. For example a

typical deaf child of age five has a vocabulary of only 25 to 50 words (200 is very unusual), whereas a hearing child may have a vocabulary of some 2000 words by the same age (Meadow, 1980). Because the profoundly deaf child has little knowledge of speech, the acquisition of lip-reading skills is diminished, which in turn affects reading and writing skills (Meadow, 1980).

1.3 Possible Solutions - Introduction

There has been a great deal of research in the area of communication aids. An attractive solution to the problem of providing acoustic information to the deaf individual is to 'replace' the auditory system with another sense, such as touch (tactile) or vision. The idea behind sensory substitutions is that the parameters of the speech waveform which are important for understanding speech and/or for learning to speak, are to be extracted and presented to the deaf person through the tactile or visual sense. A third type of system called the cochlear electrode implant is used in certain types of profound deafness where the auditory nerve remains functional but the hair cells in the cochlea, which translate the physical vibrations of sound into electrical nerve impulses, are damaged.

The choice of one of the three methods of speech information feedback depends on a number of parameters and is often the subject of heated debate among researchers in the various fields. Factors which must be taken into account when

selecting a display method include: information capacity of the sensory channel, portability of the resulting electronics (size, power consumption), appearance of the aid, comfort, ease of use including training and fitting and finally cost of the device. The user's requirements and personal characteristics (type of deafness, language ability, motivation, age etc.) also play an important role in selecting the best sensory system for feedback of speech information. A brief outline of the characteristics of the visual and cochlear implant display schemes will emphasize the advantages and drawbacks of each.

1.3.1 Visual Displays

Visual display methods of speech parameters are commonly used in schools for deaf children. With one notable exception (Upton, 1968), visual systems are non-wearable and used mainly for speech production training. The use of visual displays as an aid to lip-reading is limited because the additional visual information tends to distract the lip-reader, making it difficult to concentrate on lip and jaw movements.

The device developed by Upton, however, employed a novel display scheme in which the visual information appeared to the user superimposed on the speaker's face. Thus the deaf person could interpret the simple display while lip-reading. The performance of the Upton system was severely limited by the speech processing methods employed. In addition to the

inaccuracies in speech information extraction, the visual display limited the use of the aid to situations where the speaker and user were facing each other and the proper distance apart for the image to be superimposed on the speaker's face. The advancement in speech processing methods coupled with the significant breakthroughs in head-up displays (Scientific American, 1986), may produce a resurgence in devices similar to the Upton system in the near future.

Visual displays used for speech training present information about the correct sound and the student tries to produce the same sound by matching the pattern of their speech to that of the teacher (Houde and Levitt, 1980). Some types of information displayed include: clusters of vowel sounds arranged spatially on the screen (Povel and Wansink, 1986), vocal tract shape derived from LPC analysis (Aguilera et al., 1986; Pardo, 1982), frequency-amplitude graphs (Boyd et al., 1986), formant frequency and amplitude (Houde and Levitt, 1980) and three-dimensional displays of frequency, amplitude and time, called spectrograms (Stark, 1971).

Visual display systems can be used to provide complex information, such as the spectrogram, of the speech produced by a deaf individual. The spectrogram-type display is useful to the teacher for analyzing the difference between the distortions found in deaf speech and desired speech patterns. Complex visual information, such as the spectrogram, has also been employed as a training aid. However, the intricacy of

the display makes rapid interpretation difficult. Consequently, display systems used for training have incorporated only one or two features of the speech waveform. (For an overview of speech training systems employing visual and other feedback means see Strong, 1975. The reader is also referred to the appropriate subsection of the Bibliography section for further references.)

One advantage of visual display systems is that an element of fun may be injected in the training through the use of colourful computer graphics. This is particularly important for attracting and keeping the attention of children. Interesting computer graphic displays are already proving to be successful (Bernstein et al. 1986, Shigenaga and Hirotaka 1986, Pickett 1982). Although usually not feasible for use in a wearable aid for the deaf person, visual display systems are very useful in speech training and may also prove beneficial in teaching people to use other feedback systems.

1.3.2 Cochlear Implants

In some types of deafness, the auditory nerve remains functional but the hair cells in the cochlea, which translate physical vibrations of sound into electrical nerve impulses, are damaged. A cochlear implant bypasses the damaged hair cells via an implanted electrode, stimulating the auditory nerve directly. Some schemes use an array of electrodes in order to increase the spectral content of the information

imparted. Most implant candidates are profoundly deaf adults who are post-lingually deaf. These patients report an awareness of environmental sounds, relief from the isolation of being deaf and improved lip-reading ability.

The interest in cochlear implant research is very strong and reflects the potential gains that may be offered by cochlear implants of the future. (The reader is directed to the Bibliography section for a listing of important research in this area.) Despite the few 'star' implant patients who have had great success in using their implants, there are a number of disadvantages to the cochlear implant system, particularly for children. First the implant solution is only viable if the auditory nerve is functional. In some types of deafness, such as those caused by disease, the auditory nerve is damaged thus making a cochlear implant unsuitable.

Second, the surgical procedure and electrical stimulation usually destroy any residual hearing (Pickett and McFarland, 1985). Third, the electrical stimulation may also promote bone growth in the cochlea which increases the amount of stimulation signal required. Also associated with the surgical procedure is the risk of extra- and intracochlear infection, which although small, is certainly not inconsequential (Leob, 1985). Other potential risks include: facial nerve damage/paralysis, intracochlear scarring and fibrous tissue growth and possible vestibular complications as well as, exacerbation of tinnitus (Resnick, et al., 1986).

In addition, the surgical procedure itself has become very expensive.

All of the above factors apply in the case of profoundly deaf children, but there are additional and perhaps even stronger reasons to limit the use of implants in children. When evaluating the hearing impairment of children it can be very difficult to obtain an accurate estimate of a child's residual hearing before the age of four or five (Simmons, 1985). In many cases a child may turn out to have sufficient hearing to do as well with a traditional hearing aid, either alone or in combination with a tactile device, as with a cochlear implant at the present stage of development. Much work needs to be done in all areas of hearing aid research, including cochlear implants and at this stage in the development process cochlear implant performance has yet to consistently exceed non-invasive techniques such as vibrotactile aids (Ifukube, 1989, Carney, 1986; Tyler et al., 1986; Goldstein and Proctor, 1985). There are also the strategic problems of dealing with a prelingually deaf child with whom communication is difficult. The trauma involved with surgery, testing and training may prove very difficult for the child to manage.

Although cochlear implants may be of benefit to postlingually deaf persons and will no doubt be the solution of the future for many types of deafness, it is the view of this author and others (Resnick et al. 1986; Goldstein and Proctor, 1985; Simmons, 1982) that given the risks,

performance not exceeding current tactile devices and surgical expense; cochlear implants are not suitable for profoundly deaf children. For an opposing view see Eisenberg et. al., (1983). The unsuitability of cochlear implants for children and their expense for all deaf persons has led to a great deal of research in the area of tactile prostheses, both in the academic and the private sector. From this research the design requirements for tactile aids to speech reception and speech training have been established.

1.4 Design Requirements for a Tactile Aid

Since the work of R.H. Gault (1924) in the 1920's there has been a great deal of research in the area of tactile prostheses for the profoundly deaf. Chapter 2 of this thesis examines some of this important work. In designing a speechreading/speech training aid, there are many important criteria which must be met. The device must provide useful information about the speech waveform in a manner which is easily interpreted in conjunction with lipreading. This implies that the speech processing must be performed keeping in mind the needs of the speechreader and those of someone learning to speak. Research relating to the effects of profound deafness on speech reception and production provide guidelines for the speech processing requirements of a tactile aid. In addition to the fundamental requirement of providing appropriate information to the deaf user, there are

a number of practical issues which must be addressed in order to ensure the success of the device.

1.4.1 Speech Processing Considerations - Speech Production

Without the auditory system for feedback, speaking becomes a laborious task for the hearing impaired individual. Studies examining the speech of hearing impaired persons have illustrated the effects of the missing auditory feedback cues. Although it is recognized that no two persons, hearing impaired or otherwise, will speak identically; research has demonstrated that a number of specific differences exist between 'deaf speech' and that produced by hearing speakers, in general. Convenient categories of these deficiencies may be labelled: i) Voice Quality, ii) Rhythm, Timing and Intonation and iii) Errors of Articulation.

i) Voice Quality

The quality of speech produced by hearing impaired speakers varies in a number of ways from the speech of the hearing population. The volume of speech is usually louder (Levitt, 1973; Nickerson, 1975; Carhart, 1970; Martony, 1968) but can sometimes be quieter, particularly if the hearing loss is of the conduction type. Because the deaf person hears his/her own voice via bone conduction, but does not hear the other speaker's voice as well, the deaf speaker will lower his/her voice to match the perceived loudness of the other speaker. (Miller, 1968). The loudness of speech may also vary

erratically in the case of deaf speech (Nickerson and Stevens, 1973). Many deaf speakers have poor control over the velum, which governs airflow through the nasal cavity, resulting in speech with a nasal quality (Stevens et al., 1974; Nickerson, 1975; Colton and Cooker, 1968). In terms of the overall characteristics of deaf speech, 15 teachers of the deaf selected the following adjectives as best describing deaf speech: "flat", "tense", "throaty", "breathy" and "harsh" (Calvert, 1967).

ii) Rhythm, Timing and Intonation

Some of the adjectives selected by the teachers in Calvert's study may be the result of the lack of prosody and errors of intonation and timing which commonly occur in the speech of the hearing impaired (Nickerson, 1975). The lack of prosody or rhythm produces speech which may be described as "flat" or "tense". Timing errors contribute to the "harshness" that deaf speech often possesses. The temporal aspects of the speech of the deaf may differ, in the aggregate, from that of hearing speakers in a number of ways. Firstly, the rate of speech tends to be slower for deaf speakers (Nickerson, 1975; Boone, 1966; Nickerson et al., 1974; Colton and Cooker, 1968). Also there may be more pauses, pauses of longer duration and/or pauses inserted at inappropriate places in speech produced by deaf speakers (Nickerson et al., 1974).

Deaf speakers may also fail to make the difference in duration of stressed and unstressed sounds sufficiently large (Angelocci, 1962). Failure to indicate stressed sounds may also be confounded by insufficient variance of fundamental frequency (pitch) of the speaker's voice. In the speech of hearing talkers, the pitch typically varies over a range of 1 to 1.5 octaves (Fairbanks, 1940), with stressed syllables usually indicated by a pitch change. A deaf speaker may fail to both shorten the duration and lower the pitch of an unstressed vowel. Deaf speakers also tend to have a high average pitch (Martony, 1968; Angelocci et al., 1964). Intonation (pitch change) problems are also common. These consist of either too little pitch change, resulting in a monotone voice, or erratic and/or excessive intonation patterns (Levitt, 1972; Nickerson, 1975).

iii) Errors of Articulation

Information on the last group of problems, errors of articulation, is comprised mainly of studies done with small numbers of deaf speakers and, thus tends to be fragmented. However, there are a variety of types of articulatory problems which are most frequent. Deaf speech has been found to be quite variable, i.e. the same sound, a vowel for instance, may be spoken differently over a period of time (Osberger and Neely, 1985). This variability has been linked to the absence of acoustic feedback information (Zimmerman and Rettaliata, 1981). As well as variability in production, often there will be a substitution of one sound for another

(Stevens et al. 1974; Nickerson, 1975), failure to develop certain sounds, such as /s/ (as in sun) and /ʃ/ (as in shoe), or failure to distinguish between voiced and unvoiced sounds.

Some of the articulatory errors, such as the failure to develop the /s/ and /ʃ/ sounds are thought to be related to the fact that the articulatory gestures involved in producing these sounds are very difficult to lipread (Nickerson, 1975). In fact, Levitt (1974), has commented that the errors in articulation of deaf speakers (Smith, 1973) are similar to the patterns of confusions among phonemes in studies of lipreading performance (Erber, 1974). As well, Norber (1967) found that the sounds which are most visible were consistently produced more accurately by hearing impaired speakers, thus suggesting the important influence of lipreading on speech production.

Most of the problems associated with the speech of the hearing impaired are interrelated in many ways. Due to the interactions of the various difficulties deaf speakers encounter, it is likely that a problem-by-problem approach to training will have limited success (Nickerson, 1975). Nickerson (1975) suggests that the solution to training deaf children to speak will be in the form of a wearable visual or tactile display of his or her own and other people's speech. Such a display would employ speech features that are as rich and informative as a hearing individual can acquire via the ear.

1.4.2 Speech Processing Considerations - Speechreading

The high degree of correlation between sounds difficult to produce for deaf speakers and those sounds which prove challenging to lipread provides encouragement that the device may be able to assist in both speech production training and in speechreading. As an aid to lipreading, the main goal is to provide cues for sounds which are identical visually but not auditorily as well as for speech produced via articulators that are not visible. The device may also impart information about those sounds which are visible in order to provide reinforcement to children learning to lipread and for use in situations where the view of the speaker is less than ideal.

A great deal of research has been undertaken on the relationship between speechreading performance and the various factors affecting accuracy in both hearing and hearing impaired subjects. Speechreading is a difficult task because a number of speech elements look alike, even under ideal conditions of view. Therefore the observer must also extract meaning from the language context and/or situation. The difficulties of lipreading are further confounded by the fact that some articulatory movements occur deep within the oral cavity and hence are not visible, well still other speech movements have decreased visibility due to coarticulation effects (Erber, 1973).

In terms of the difficulties in visual recognition of specific speech sounds, most research has been done at the phoneme level. (The reader is referred to Appendix I, for a list of the approximately 40 phonemes and corresponding symbols, of the English language.) Vowels, in general, are easier to lipread than consonants (Markides, 1979). According to studies done by the American Society for the Hard of Hearing (1943), the most difficult consonants to identify visually are those where the place of articulation is not visible, such as /k,g,ŋ and h/, with the easiest being those which have a place of articulation at the front of the vocal tract (labial) such as /b,p,m,f,v,w/ (see Appendix I for phoneme classes). The most difficult vowels were: /ɛ,ɪ,ʊ,ə,ʊ,ə / with /a,æ,ɔɪ,aʊ/ among the easiest. Note however, that this classification of difficulty ignores the fact that many of the phonemes have identical visual appearances, i.e. /b/ and /p/ are difficult to discriminate.

Profoundly hearing impaired children, for instance, were found to have the highest confusions in visual recognition of consonants for those sounds that differed in the manner of production, but appeared similar visually (Erber, 1979). The place of articulation, which can be lipread, is not sufficient to identify most speech sounds; for this, the additional information provided by the manner in which the sound is produced (manner of articulation) and/or the voiced-voiceless distinction is required. Consonant sounds may be classified as fricatives, affricates, nasals, plosives and

approximates according to the manner of articulation (National Academy of Sciences, 1971). An example is the consonant group /t,n,d/ which all share the same place of articulation and thus appear similar via lipreading, but differ in manner of articulation and/or voicing. (/t/ is voiceless, stop, /n/ is voiced, nasal, /d/ is voiced, stop.) Other research has also found that the majority of speechreading errors occur in the manner of articulation and voiced-unvoiced decisions (Binne et al., 1976; Hack et al., 1982; Erber, 1973; Walden et al, 1977).

Groups of sounds that appear similar visually such as /t, n, d/ are often termed visemes (Fisher, 1968) and are the visual equivalent to phonemes. Research has shown that the 40 phonemes of the English language are mapped to only four stable viseme groups, when speech is viewed under normal conditions (Jeffers and Barley, 1977). In an ideal speechreading environment, the number of distinct visual groupings increased to nine and an additional five groups were found to be present under certain conditions of view (Jeffers and Barley, 1977). The phoneme groupings in the four visemes observed during usual conditions by Jeffers and Barley (1977), are shown in Table 2. Other researchers have found similar viseme categories, with the number varying between four and sixteen, depending on the experimental setup and procedure (Erber, 1973; Walden et al., 1977; Binnie et al, 1976; Woodward and Barber, 1960).

Phonemes	Viseme - Articulator Shape
/ f, v /	Lower lip to teeth
/ w, r, u, ɔɪ, o, ʒ /	Lips puckered- narrow opening
/ p, b, m /	Lips together
/ ə, ð /	Tongue between the teeth
/ others /	No stable grouping

Table 2 Visemes and Corresponding Phonemes (Jeffers and Barley, 1977)

The limitations of speechreading become clear when one realizes that there are only 4-16 distinctive visually perceivable features to represent the 40 phonemes of the English language. There is a further complication, in that the most visible speech sounds are unfortunately, not the most frequently spoken. The /f,v/ category is one of the more stable visemes, but these sounds only occur approximately 3.15% of the time (Brooks, 1984). Conversely, four of the most frequently used sounds, /t,d,n,l/, account for 21.10% of phonemes in conversational speech, but are within the same viseme group and therefore are not distinguishable via lipreading (Brooks, 1984).

Thus, although place of articulation can be derived from speechreading, this is often not sufficient for the perception of many speech sounds. The manner of articulation

and/or voicing information are required in order to differentiate sounds within the same viseme category. A tactile aid to lipreading therefore, should attempt to provide this missing information. Since many speech sounds contain different spectral energy peaks, due to closures and resonances of the vocal tract, spectral energies provided tactually may also aid in the speechreading process. As mentioned earlier, the intonation patterns of speech are often important to the meaning of the message, but are not available to the lipreader. Chapter 2 will examine some speech processing strategies of past tactile aid research and their performance in supplementing the lipreading task.

1.4.3 Other Design Issues

In addition to the obvious requirements involving speech processing, there are also a number of practical design considerations which must be addressed. First among these are the information bearing properties of the tactile sensory channel and its limitations. It is important that the speech information not exceed the transmission capacity of the sense of touch. Section 2.2 outlines the capabilities and limitations of the tactile sense and the implications in the design and performance of a tactile aid for the deaf.

Just as it takes a hearing child time to learn to use the articulatory system for speech production, the deaf person will require time to learn how to use the tactile/visual information to interpret the speech of others

and as feedback for his/her own speech. Thus to be successful, the tactile system must be used regularly and therefore be readily available to the user, i.e. not tied to a laboratory. Therefore the system must provide real-time conversion of sounds to tactile stimulation and be compact (portable).

Ideally a wearable version should be evolved from the portable system. In order to gain acceptance among the deaf population, the wearable version should be small, lightweight, and inconspicuous. The portable system should also have low power consumption. Any system not meeting the portability conditions, will not likely receive a thorough evaluation, due to its confinement in the laboratory. Systems that do not evolve to wearable versions will not receive widespread use among the profoundly deaf population, despite the potential gains in communication ability.

Chapter 2

2.0 Introduction

Given the problems associated with profound deafness, as discussed in the previous chapter, there has been a great deal of interest in the area of speech reception aids, employing the tactile sense. Section 2.1 explores some of the previous work done in this area. The limiting factor in tactile aids is most often the information transmission capacity of the tactile sensory channel. The characteristics of the tactile channel are discussed in Section 2.2. Based on the problems associated with profound deafness, previous research on tactile aids and the characteristics of the tactile sensory channel, principal components analysis (PCA) of speech spectra for use in a tactile aid, is proposed in Section 2.3.

2.1 Survey of Tactile Aid Research

The development of tactile sensory aids dates back to the 1920's, with much of the pioneering work done by R.H. Gault (Gault, 1924, 1934). Gault's system involved amplification of the speech signal which was used to drive a vibrating unit. This device provided information about the rhythm, duration, stress and relative intensity of the speech signal. Gault's experiments with 22 deaf subjects yielded improvements in lipreading ability of 10-15%, when lipreading was supplemented with the tactile device.

This early research demonstrated that deaf children and adults could benefit from the tactile presentation of speech information. The gains however, were not as large as Gault had hoped and the interest in tactile aids declined. The invention of the transistor led to the development of wearable hearing aids, which provided a great breakthrough for a large portion of the hearing impaired population. Unfortunately for those individuals with profound hearing losses, for whom amplification aids were of little use, the interest in tactile aid development further decreased with the advent of compact auditory hearing aids.

Eventually the inadequacy of auditory hearing aids for the profoundly deaf revived interest in the development of tactile aids in the 1950's, and research has continued to the present time. But only recently (previous decade) have the advances in VLSI (Very Large Scale Integration) electronics, coupled with the advent of efficient vibro- and/or electrotactile transducers allowed the design of wearable tactile sensory aids. There are a number of excellent overview papers available detailing both laboratory and wearable tactile systems (Bernstein, 1988; Kirman, 1973; Levitt, 1973; Pickett and McFarland, 1985; Pickett, 1963; Proctor and Goldstein, 1988; 1985; Sherrick, 1984; Weisenberger and Miller, 1987; and Weisenberger, 1988a).

Tactile aids may be categorized with respect to either the speech processing schemes or by means of the type of output display and transducer design. Various output methods

and tactile channel characteristics are discussed in Section 2.2. In the review and analysis to follow, tactile aids will be divided into two broad classes, those which extract a single speech characteristic, and others that employ multiple speech features. The scope of this thesis does not allow a detailed treatment of the entire body of previous research, but the most promising systems will be discussed. (The reader is referred to the appropriate section of the Bibliography for further references.)

2.1.1 Single Speech Feature Tactile Aids

Single feature systems were the first to become wearable, and thus the first available commercially. The wearability allows deaf children to use single channel devices extensively during the years when auditory communication skills normally develop. Single feature systems are designed principally as aids to lipreading; hence they provide information about speech that is normally not available visually. Three speech features commonly used in the single channel devices are amplitude envelope, frequency content, and pitch (fundamental frequency).

Systems that calculate the amplitude envelope of the speech signal provide the wearer with information about the timing, rhythm and duration of speech which is very difficult to derive via lipreading, as mentioned in Section 1.3.2. Prosody information also tends to be absent in the speech of profoundly deaf children. Information about the intensity,

also lacking in deaf speech, is provided by the speech envelope tactile aids.

Amplitude envelope systems typically employ amplitude modulation of a fixed frequency carrier signal to drive a vibrotactile transducer worn on the wrist (United States Patent #4,728,934, 1988) or on the sternum. Two envelope type devices are available commercially. The first marketed as the Minivib 3, is sold by Special Instruments AB, Stockholm, Sweden and the second, is available from the Audiological Engineering Corp. of Boston, as the Tactaid I.

The Minivib 3, is based on the work of Spens and Plant (1983). Spens and Plant found that use of their device increased lipreading scores in profoundly deaf adults. Weisenberger and Russell (1986) also tested the Minivib device. They artificially deafened normally hearing subjects during testing, and had them perform sound detection, phoneme discrimination, and syllable counting tasks without the aid of lipreading. Weisenberger and Russell found the Minivib system to be effective for sound detection and allowed the subjects to discriminate phonemes 25% above the chance level. The Minivib 3 has been chosen by Proctor and her colleagues for use in a study with deaf preschool age children, due to its compact size, light weight, commercial availability, relatively low cost (approx. \$450 U.S.), and overall performance (Proctor, 1988, Proctor and Bickley, 1988).

The Tactaid I, from Audiological Engineering, is very similar to the Minivib device, although somewhat more

expensive (approx \$650 U.S.). Users of this device report improved lipreading accuracy, more understandable speech production, awareness of sounds in their environment and general relief from the isolation of profound deafness (Audiological Eng., 1986). This device and a similar system developed at Johns Hopkins University have been evaluated extensively by Proctor and Goldstein (1987, 1985). Studies with four profoundly deaf children indicated that not only did the language performance of the children increase with time, but the profoundly deaf tactile aid users exhibited improvement in language comprehension at a rate greater than that shown by hearing children of the same age (Proctor and Goldstein, 1987). Other tests demonstrated that the expressive spoken language abilities of the children also increased substantially.

A third commercial device, the Minifonator from Siemens Hearing Instruments, combines speech envelope extraction with information about the peak frequency band of the speech signal. The instantaneous amplitude of the envelope is used to control the vibrator carrier amplitude, while the variation in frequency of the spectrum peak, produces a frequency modulation of the carrier signal. In this manner, two types of information are coded to a single vibrator. Originally designed as an aid to speech instruction for deaf-blind children, the Minifonator has been shown to provide information on the voice/voiceless distinction, long/short, aspirated/unaspirated and plosive/fricative /rolled

characteristics of speech to deaf subjects (Carney and Beachler, 1986; Schulte, 1977). Weisenberger and Russell, (1986), compared the Minifonator device with the simpler Minivib 3 and found nearly similar performance in most tests, despite the addition of frequency modulation to the carrier signal. This may have been due to the limited range of vibrotactile frequency perception.

The last category of single channel tactile aids attempt to extract the pitch (fundamental frequency) of the speech signal and provide information about its frequency to the user. This may be accomplished by a straightforward frequency to amplitude conversion to amplitude modulate the vibrator signal, or may involve a multiple vibrator display. Experiments using auditorily presented pitch cues to supplement lipreading for hearing subjects have demonstrated the utility of voice pitch as an aid to speechreading (Breeuwer and Plomp, 1986; Grant et al., 1985).

Grant and his colleagues found that normal-hearing subjects improved their scores in the tracking task (De Filippo and Scott, 1978), from 41.1 words per minute, with speechreading alone, to 73.6 words per minute when pitch information only was presented auditorily to hearing subjects as a supplement to lipreading. Breeuwer and Plomp also discovered voice pitch to be an effective aid to speechreading. Recall also, from Section 1.3.1, that deaf speakers often have difficulty with pitch changes (intonation) and sometimes have a higher than normal average

pitch. Thus a device which provides feedback about voice pitch may also serve as a speech production training aid for profoundly deaf children.

Although there have been no published results of tactile fundamental frequency display devices used as speech training aids, a number of groups have demonstrated that tactile pitch systems provide an effective supplement to speechreading. Boothroyd's group at the City University of New York developed and evaluated a number of single channel tactile pitch displays as well as multichannel systems (Boothroyd et al., 1988; Boothroyd and Hnath-Chisolm, 1987; Boothroyd, 1985; Boothroyd et al., 1986a; 1986b; Boothroyd et al., 1985; Boothroyd, 1984). They found both types of displays enhanced perception of voicing and stress contrasts of speech, but the multichannel system proved to be more effective in transmitting intonation information (Boothroyd et al., 1988). Grant et al. (1986), reported that patterns of intonation, stress and phrase structure, all of which are not easily lipread, were perceived by users of their 10 electrode pitch display. Grant's group used an electrotactile array, where each electrode represented a narrow bandwidth of pitch frequency within the total range of 100-300 Hz for female speakers and 57-200 for male speakers. Two subjects learned to recognize whether a sentence was a question or statement, to identify the location of a stressed word in a continuously voiced sentence and to identify the location of a phrase boundary; all at an accuracy of >90% without speechreading.

When speechreading was supplemented by use of the tactile aid, scores in these tasks approached 100%, as compared to near 50% for lipreading alone. Tests with a profoundly deaf subject using the tracking procedure for training and testing demonstrated that tactile and visual information were complementary and together allowed speech reception at speaking rates and conditions approaching normal communication settings.

Recently, Bernstein and her colleagues studied six transformations of speech fundamental frequency to vibrotactile stimulus (Bernstein et al., 1989). The transformations tested were: i) feedthrough - which generated a tactile pulse for each glottal pulse, ii) fixed frequency - provided a 250-pps (pulses per second) signal during voiced segments of speech, iii) scaled F0 - tactile signal period was a scaled function of pitch period, iv) log scaled F0 - tactile period was a logarithm of the pitch period divided by the mean pitch, v) difference scaled transformation - tactile pitch was a function of present pitch divided by a weighted sum of previous pitch values and vi) the amplitude modulated dual fixed frequency transformation employed two vibrating devices, one at a frequency of 40 Hz, the other @ 400 Hz, setup so that increasing pitch frequency caused the amplitude of the 40 Hz vibrator to decrease and the 400 Hz vibrator amplitude to decrease and vice-versa for decreasing pitch frequency. Bernstein's group reported that the pitch information provided a small, but significant improvement in

intonation identification and a significant improvement in stress identification, but there was little difference in performance between the six transformations. Comparisons of the lipreading alone, tactile alone and lipreading plus tactile modes for the intonation tasks suggested that the users were not taking full advantage of the tactile devices when lipreading since the scores for the "tactile only" case were higher than the tactile plus speechreading category. Bernstein et al. suggest that the complex signal resulting from the transformations that map voice pitch to a single vibrator may be difficult to interpret without extensive training.

Indeed, the most successful of the single feature tactile aids, employed very straightforward mapping of the selected speech feature to the tactile sense. In most cases the transformation resulted in a fixed frequency, amplitude modulated signal. Section 2.2, dealing with the characteristics of the tactile sensory channel will explore this situation further. Studies of the performance of the single feature tactile aids have demonstrated that the amplitude envelope and pitch information can be used as successful supplements to speechreading. In trials conducted with deaf children, Proctor and Goldstein (1987) found that tactually presented amplitude envelope information did aid the children in acquiring lipreading skills. However, tests done later without the aids, resulted in scores similar to those achieved with the aid. It seemed that the single

channel systems did not act as a sensory substitute for the auditory system, but did aid in the development of lipreading skills. Although single channel devices can be beneficial aids to speechreading, multifeature tactile systems may provide equal or better gains in language comprehension, while having better potential as a sensory substitute; i.e. provide more information to the user.

2.1.2 Multifeature Tactile Aids

Single feature tactile aids provide information about the temporal changes in a speech signal, whether it be the amplitude envelope or the pitch period display system. In order to better model the function of the human ear, many researchers have attempted to provide more detailed information about the spectral aspects of the speech signal. This information is transmitted via an array of tactile transducers. By supplying information about more than one frequency band it was hoped that superior performance over the single feature devices would be realized. Three strategies have been used in multifeature systems: i) two or more bandpass filters used to detect the energy in certain frequency bands, ii) tracking of the resonant frequencies (formants) of the vocal tract and iii) frequency content information presented in terms of temporal changes.

1) Bandpass Filter Methods

The speech spectrum display methods range in complexity from simple dual bandpass filter systems to sophisticated 16 or 32 channel banks of bandpass filters, called vocoders (voice-coders). At the low-complexity end of the scale are two channel systems that provide information about the energy in only two spectral bands. Breeuwer and Plomp (1984), reported that information about sound pressure levels in two appropriately selected frequency bands, presented auditorily to hearing subjects, improved speechreading scores on short sentence tests. Using sound pressure information from two 1-octave bandwidth, bandpass filters, with centre frequencies of 500 and 3150 Hz, subjects were able to achieve scores of 87% correct vs. only 22% by lipreading alone. Recently a tactile aid, with similar speech processing characteristics has become commercially available (Tactaid II from Audiological Engineering Corp., 1988), but to date no performance data are available.

The first of the vocoder type systems, was developed by Guelke and Huyssen (1956) and used eight filters over the frequency range of 400 to 2880 Hz. Although their test data were limited, Guelke and Huyssen were aware of the sharp amplitude sensitivity as a function of frequency for the tactile sense and bypassed this problem by using the spatial location of vibration to represent frequency. This scheme has been used since in research efforts on tactile vocoders. The general configuration is as shown in Figure 3. The

amplitude of the energy in each frequency band is used to amplitude modulate a fixed frequency carrier signal, with the frequency chosen to provide optimum tactile amplitude sensitivity. The vibrators are then placed at different locations on the body.

Figure 3 Tactile Vocoder Configuration

Pickett and Pickett (1963), used a similar configuration, as illustrated in Figure 3, with 10 vibrators applied to the fingers of the user. They performed evaluations of the system without the aid of speechreading and found identification and discrimination of vowels to be

good overall. This seemed to be related to the formant structure of the sounds. Consonant discrimination test results ranged from 99.5% for the /s-t/ pair to 22% for the /p-b/ pair.

Saunders and his group applied their expertise in electrotactile stimulation, (gained through the design of aids for the blind), to a 20 channel electrotactile vocoder (Saunders et al., 1976). The user of this system wore a linear array of 20 electrodes, spaced 1/2 inch apart, on a belt around the abdomen. Initial tests revealed that subjects could learn to correctly identify words within simple sentences using the tactile aid alone. Franklin and Saunders (1984), expanded the 20 channel device to 32 channels and dubbed it the Teletactor. Training of three profoundly deaf children, over a two year period, resulted in improved speech intelligibility comprised of improved rhythm, reduction in nasality and voice tension and reduction in voice/voiceless confusions among consonants. In terms of speech reception, six deaf children demonstrated improved speech comprehension when using the device. However, comparisons to a 16 channel electrotactile system, the Tacticom, revealed that the Tacticom equalled the performance of the larger 32 channel system. Although the number of bandpass filters was halved in the Tacticom, the bandwidths and centre frequencies more closely approximated those found in the human auditory system (Franklin and Saunders, 1984).

Indeed, the most successful of the vocoder devices to date employs a bank of bandpass filters, designed to model human audition. This system has been developed by an interdisciplinary team at Queen's University (Mason et al., 1981) and employs a filter bank designed by Searle et al., (1978; 1979) which was in turn based on the auditory models of Kiang and Moxon (1974). Brooks performed an extensive evaluation of the 16 channel vibrotactile aid (Brooks et al., 1987; Brooks et al., 1985; Brooks, 1984; Brooks and Frost, 1983). Brooks and Frost (1983) tested the system in the tactile only mode and found that two normal-hearing subjects learned to identify 70 and 150 words correctly with 40.5 and 55 hours respectively, of training. It was also noted that many identifications not possible in the first 10 - 20 hours of training could be learned with more training. In fact one subject eventually acquired a 250-word vocabulary through the tactile vocoder with a total of 80.5 hours of training (Brooks et al., 1985). Further testing with two profoundly deaf teenagers demonstrated that deaf subjects could also learn to identify words in the tactile mode alone. In addition, the subjects could discriminate between voiced and unvoiced examples of stop and fricative consonants with 84.5% accuracy.

The Queen's tactile vocoder was also found to be effective as a supplement to lipreading, with performance improvements of 25-29%, over speechreading alone on three different speech comprehension tasks (Brooks, 1984). Although

the performance data was acquired using a laboratory vocoder system, the feasibility of wearable tactile vocoders has been demonstrated by Katz and Wentzel (1979), Engebretson and O'Connell (1986) and Engebretson et al. (1986).

The final area of tactile vocoder research differs from the systems described previously in that a 2-dimensional array of tactile transducers is used. The MESA, multipoint electrotactile speech aid, developed by Sparks et al. (1979) used a 36 X 8 array of electrodes mounted on a belt. Each of the 36 columns was assigned to one of 36 bandpass filters which spanned the frequency range of 85-10,000 Hz. The 8 rows represented amplitude steps of 5 dB per row, thus the filter energies were transformed to the spatial domain of the tactile sense. Sparks et al (1979) found that after an initial learning period, the MESA device combined with lipreading allowed improved performance over lipreading alone. As in Proctor and Goldstein's study, after extensive learning, performance in lipreading plus the MESA and lipreading alone were nearly similar. Again, the tactile aid facilitated the acquisition of lipreading skills, but did not perform as a sensory substitute to hearing.

Yeni-Komshian and Goldstein (1977), used a similar mapping of amplitude and frequency to a 2-D vibrotactile array. In their system, 18 columns of the 24 x 6 array from the Optacon reading aid for the blind (Bliss and Linvill, 1967) was used. The amplitude of each filter output was coded as spatial position, as in the MESA device, but rather than

limiting the stimulation to one vibrator per column, or filter, the Optacon display operated in a bar graph mode with all vibrators in a column turned on for full amplitude output from the frequency channel. Reported performance data indicate that subjects could learn to discriminate three vowels, four vowel durations and four spondee words, but learning to identify the vibrotactile patterns required many hours of training.

In yet another variation on the vocoder scheme, Ifukube (1982) has developed a 16 channel tactile vocoder that outputs to a miniature 16 X 3 piezoelectric vibrotactile array worn on one fingertip. Ifukube's vibrator array consists of 48 piezoelectric reeds, each of which moves a pin to contact the pad of the finger. The array is organized in 3 columns of 16 pins, with a spacing of 3 mm between rows and 1 mm between adjacent pins in the columns. The outputs of the bandpass filters were scanned across the three rows, thereby allowing the user to perceive the present spectrum as well as the two previous spectrums simultaneously. It was hoped that such a system would overcome the poor temporal sensitivity of the tactile sense to rapid changes. However, the complex stimulus generated by the scanned display proved to be very difficult to interpret.

In the second iteration of the system, an additional 3 x 2 array was designed to provide cues as to the type of speech sound presented. In this embodiment, the 16 x 3 array was setup to output using all three rows, for each channel

simultaneously and thus was essentially performing as a 16 x 1 array. However, in parallel to this display, a second array was worn on the thumb and indicated the type of speech sound. The six sound classes used were, semivowel, nasal, fricative, plosive, voiced and unvoiced. The limited test results available indicate that the cues may be an effective supplement to the vocoder, but the accuracy of the cue extraction software was poor with connected speech. Although there are no performance data of Ifukube's cued vocoder as an aid for the deaf, the compact, low power vibrotactile display method is an important breakthrough; despite the disadvantage of dedicating the use of one hand to speech interpretation.

There are many configurations of the vocoder-type tactile aids for the profoundly deaf. Differences exist in the number of filters, frequency range of analysis, bandwidth and filter centre frequencies and output display means. All of the vocoder designs do however, share a common goal and that is to provide the user with an indication of the frequency content of speech. The success of the Queen's vocoder indicates the utility of such information, both as a supplement to lipreading and for the perception of words and phonemes when lipreading is not possible. The main disadvantage of the vibrotactile vocoders is that the large number of transducers required limits the wearability of such systems due to high power consumption. Ifukube's output device provides one solution to the problem, but at the expense of dedicating the use of one hand to speech

interpretation. Another solution to the problem of large transducer arrays, involves providing only information about the frequency peaks, called formants in the speech spectrum.

ii) Tactile Formant Display Systems

Formant extraction is a common scheme with cochlear implant speech processors due to the popular theory of speech perception involving the frequency and amplitude of the spectral peaks in speech (Fry et al., 1962; Delattre et al., 1955; O'Connor et al., 1957) and the use of formants in speech synthesis (Klatt, 1987). One tactile implementation of the formant display uses the speech processor of the cochlear implant system produced by Nucleus Limited, of Melbourne, Australia (Blamey and Clark, 1985). This processor selects the transducer to be stimulated, based on the largest spectral peak between 800 and 4000 Hz, which corresponds to the second formant of vowel sounds. The pulse rate of stimulation is proportional to the pitch of the speech and the pulse width (perceived as intensity or loudness) is a compressed function of the speech amplitude envelope. Thus two speech characteristics, pitch and amplitude envelope, which were found to be successful in single channel tactile aids, are combined with the second formant location in the tactile display.

The display itself is a novel electrotactile array of 8 electrodes. The 8 electrodes are secured to the inside and outside of each finger on one hand. A larger ground electrode

is worn on the underside of the wrist. All electrodes are secured with flexible straps, so as to allow normal use of the hand. The placement of the electrodes allows stimulation of the digital nerve bundles of the fingers for effective percepts, while the electrical stimulation method minimizes power consumption.

Speech discrimination tests were carried out with seven hearing subjects and four hearing impaired adults (Cowan et al., 1988; Blamey et al., 1988). For the speech tracking task the hearing subjects improved their scores from 36 words per minute (wpm) during lipreading alone, to 55 wpm when lipreading was supplemented by the tactile aid, after 70 hours of training. With 35 hours of training the hearing impaired group improved from 24 wpm, in the lipreading only mode, to 37 wpm in the tactile plus lipreading mode. Extensive evaluations using a variety of speech comprehension tests, demonstrated the electrotactile unit was beneficial as a supplement to lipreading and provided information which aided the recognition of both words and phonemes, as well as duration, voicing and manner of articulation speech features.

By combining pitch, amplitude envelope and second formant frequency information to drive a novel electrotactile array, Blamey and his colleagues have developed a promising tactile aid for the deaf. While performance evaluation is ongoing, initial results are very encouraging. Currently however, the speech processor and electrode interface circuitry require two enclosures and therefore make the

system somewhat cumbersome. Large electrode areas are often required to reduce the current density in order to minimize painful sensations associated with electrical stimulation of the skin, and this may preclude the use of the system by young children. Last, both the vocoder and formant extraction schemes address the speech processing problem from the spectral domain. There are other theories of speech perception that consider the importance of the temporal aspects of speech (Baum and Blumstein, 1987; Porter et al., 1987; Walsh and Diehl, 1987; Scott, 1980) and at least one research group has based their speech processor on these theories (Scott, 1981; 1980).

iii) Combined Spectral / Temporal Tactile Speech Processor

Scott's speech processing strategy differs from others in that he stresses examination of the integrated signal, that is the relationships among spectral features in the complex speech signal, rather than focusing on the individual spectral features themselves. Traditional methods of speech analysis rely on the time-to-frequency transformation of the speech spectrograph and indeed this type of frequency analysis is performed in the cochlea of the human auditory system. Scott, however, hypothesizes that the spectral separation performed by the human auditory system is done as a means of interfacing the high frequency components of the speech signal to the relatively slow responses of the neurons and that higher up in the perceptual system, the frequency

components are re-integrated in order to retain an accurate representation of the acoustic signal. This theory borrows much from the Gestalt psychologists, who studied visual patterns such as the one shown in Figure 4. One perceives this figure as a diamond, despite the fact that it is composed of squares, circles and dots. The relationship of the squares, circles and dots of Figure 4, combine to create the percept of the diamond and Scott believes this is analogous to the speech signal in which a complex combination of many sinusoids creates the percept of a speech sound.

Figure 4. Complex visual pattern (after Scott, 1980)

Continuing with this analogy, the transform of Figure 4, as illustrated in Figure 5 represents a similar effect as spectral analysis of the speech signal. In Figure 5, the components of Figure 4, that is the squares, circles and dots, are separated in much the same manner as the frequency components of speech are separated in the spectrograph. Notice that although the components of Figure 4 are more

readily displayed in Figure 5, one no longer perceives the original diamond shape.

Figure 5. Transformation of Fig. 4, to organization by size of elements comprising Fig. 4. (after Scott, 1980)

Scott believes that, in addition to the speech spectrum, the speech waveform should be examined in order to determine how the various spectral components are integrated. In order to test his hypothesis, Scott performed some interesting speech perception tests (Scott, 1976). Scott also applied his theories of speech perception to the area of tactile aids as supplements to lipreading. According to his hypothesis that the speech waveform is important in the perception of speech, he projected that a single channel tactile aid providing speech envelope information would outperform a vocoder type device, for the comprehension of sentence level material.

Cole and Scott (1974) examined the role of the speech waveform envelope in perception and hypothesized that it provides a single, slowly varying reference for the many rapid spectral changes in speech and thus aids the perceptual system in integrating information over syllable length time periods and in maintaining the temporal order of the spectral

components. Therefore the single channel tactile aid, providing the envelope information should produce better performance on sentence length speech segments, but the vocoder system may be more beneficial for short speech segments such as syllables. Tests performed by Scott (1974), using a single channel tactile system and a 5 channel tactile vocoder supported the hypothesis. In "same" or "different" tasks, subjects performed better on the sentence material when using the speech envelope aid as opposed to the vocoder and vice-versa for the syllables.

Scott concluded that a tactile speech reception aid should therefore, include an integrated amplitude envelope display, as well as spectral information. With this goal, a series of design iterations led to the development of the SRA-10 tactile aid as shown in block diagram form in Figure 6 (De Filippo, 1984; Scott, 1981; 1980).

The SRA-10 tactile aid divides the speech signal into three frequency bands. The first is a high pass filter with an 8 kHz corner frequency and corresponding envelope detector. The envelope amplitude from this high frequency 'indicator', amplitude modulates a lowpass filtered ($f_c = 250$ Hz) peak clipped noise source which drives one of three vibrators. The second frequency division is a bandpass filter centred at 2.4 kHz that amplitude modulates the same noise source and drives the other two vibrators, which are worn on either side of the first vibrator. The last channel is a

Figure 6. SRA-10 Tactile Speech Reception Aid (Scott, 1980)

bandpass filter with a pass band of 250 to 900 Hz and serves to detect the first formant of vowel sounds. The frequency range of this signal is converted to a range better perceived by the tactile sense (15-200 Hz) which is then amplitude modulated by the envelope magnitude and added to the other two signals. Thus first formant frequency and amplitude information is fed to all three transducers.

The combinations of speech features as shown in Figure 6, creates distinct patterns of sensation depending on the type of speech sound. For example, the three vibrators worn on the chest, cause an high frequency fricative, such as /s/ to be perceived as an aperiodic sensation at the centre of the chest. Aperiodic because the vibrator carrier signal is

the peak clipped noise and centre location because all of the spectral energy is concentrated in the high frequency region. Vowels are perceived as periodic sensations since the carrier signal in the lower channel is periodic. Scott states that this type of display preserves the time-integrated envelope of the speech signal because the user senses all signals in the centre of the chest, rather than in a spatial array of vibrators employed in vocoder schemes.

Scott's tactile aid has undergone tests with normal hearing subjects, (Scott, 1980; De Filippo, 1984) and with profoundly deaf children (Friel-Patti and Roeser, 1983). The hearing subjects were trained and tested using the tracking method, in both the lipreading only and tactile plus visual conditions. Scores and learning curves were superior in the aided condition for all subjects but one. The best performance was a rate of 55 wpm after 8 hours of training, four of which were done without the tactile system (Scott, 1980). In a similar test one subject achieved a score of 76 wpm, after 10 hrs. of training, 5 of which was in the unaided condition (De Filippo, 1984). This score is approximately 78% of the normal wpm score achieved by hearing subjects.

In Friel-Patti and Roeser's study, four profoundly deaf children used the SRA-10 tactile aid for 4.5 months and then continued the same speech communication training program, without the aid, for another 4.5 months. The subjects overall communication abilities were evaluated throughout the nine month period. The communication index scores, for the

children increased from 96 to 232 during the period when the children used the tactile aid. The first test session without the tactile aid produced an average score of 190, thereby demonstrating that the increase in performance with the tactile aid was not due entirely to the specialized training that the children received. It was also observed that there were no adverse reactions by the children toward using the vibrotactile aids. In addition, teachers and parents reported that the elicitation of spontaneous vocalization was much easier when the children wore the tactile aid.

Although Scott's vibrotactile aid has undergone only limited testing (6 hearing and four profoundly deaf subjects), the results are encouraging. These test results also stand to support Scott's hypothesis that the speech waveform envelope is an important cue for the perception of sentence length speech. Currently the electronics of the SRA-10 system are too large to be wearable, but the use of either custom integrated circuits or digital signal processing techniques would allow a compact version to be implemented.

2.1.3 Summary of Tactile Aid Research

In sections 2.1.1 and 2.1.2 some of the research in the area of tactile communication aids for the deaf, was presented. Tables 3a and 3b (following pages) summarize some key aspects of the aids reviewed. The reader is cautioned however, not to fall to the temptation of comparing the aids directly against one another in terms of the stated

performance. The potential for variability is exceedingly high and differences in subjects (hearing ability, age, sex, motivation, etc.), subject numbers, training methods, training time and testing differences (tasks, procedures, statistical analysis, etc.), make direct comparisons nearly impossible.

One example is the effect of the subject's hearing ability and experience on the lipreading scores. Normally hearing, inexperienced speechreaders tend to derive greater benefit from tactile aids than do deaf subjects who are experienced lipreaders. The gains in speechreading ability for the hearing subjects, in addition to the effects of the aid, may be due to their knowledge of language, the training procedures or perhaps the tactile aid providing more benefit to users with little lipreading ability. If this latter factor is the case, then is it that the aid facilitates the training of lipreading, or does it provide cues that confuse experienced lipreaders? It is possible to test for the effects of training on the performance evaluation of the tactile aid but, unfortunately, many researchers simply chose to ignore the problem.

In terms of test methods, the tasks used to evaluate the aid can be very different between groups, particularly in terms of the length of the speech stimuli. It was demonstrated by Scott (1974), that the discrimination of syllable length speech and sentences will favor different

Device Name	Reference	Transducers Type-#-Loc.	Speech Processing	Notes and Performance
Gault	Gault, 1924	Vib. - 1 - hand	Amplitude Envelope, Speech applied directly to vib.	10-15 % increase in lipreading ability.
Minivib3	Spens & Plant, 1983	Vib. - 1 - wrist or sternum	Ampl. envelope of speech, AM a 250 Hz carrier	Artif. deafened adult subj. - phoneme discrimination w/o lipreading @ 25% above chance.
Tactaid I	Proctor & Goldstein, 1987	Vib. - 1 - wrist or sternum	Ampl. envelope of speech, Ampl. Mods. a 250 Hz carrier	Four profoundly deaf children showed incr. in language comprehension and speech production.
Mini-fonator	Weisenberger & Russell, 1987	Vib. - 1 - wrist or sternum	Ampl. envel. AM & Freq. of peak ampl. FMods a 250 Hz carrier	Similar performance as Tactaid I and Minivib, despite added freq. info.
-	Grant et al., 1986	Elect. - 10 - forearm	Pitch extraction-> enables a vib. representing a narrow freq. band within 50-300 hz.	Two artif. deafened subj. increased lipreading ability by 40%.
Tacticom	Saunders, et al., 1976	Elect. - 16 - abdomen	Vocoder design (1/3 Octave filters)	3 profoundly deaf children over a 2 yr period improved both speech comp. and speech production. 16 ch device was found to outperform a 32 ch one.
Queen's Vocoder	Brooks, 1984	Vib - 16 - forearm	Vocoder based on human audition.	Artif. deafened adults improved lip-reading by 25-29%. One subj. learned a 250 word vocab. w/o lipreading in 80.5 hrs.

Table 3a Summary of Tactile Aids Reviewed

Device Name	Reference	Transducers Type-#-Loc.	Speech Processing	Notes and Performance
MESA	Sparks, et al., 1979	Elect. - 8x36 - abdomen	Vocoder, 36 filters x 8 amp. levels coded to 36x8 vibrators. Freq. range 80-10kHz.	Some improvement in lipreading ability, but display was very difficult to use.
.	Ilukube, 1982	Vib. - 16x3 & 3:2 - fingertips	Vocoder coded to 16x3 array. Speech sound class (plosive, fricative, voiced etc) coded to 3x2 array.	No performance data available. Low power (piezo-electric vibrators) display is unique, but requires the use of one hand.
Tickle Talker	Blamey, et al., 1988	Elect. - 8 + grd - sides of fingers	Frequency band of maximum energy mapped to electrode position, stimulus amplitude ~ to speech amp. envel. and pulse rate ~ pitch.	Unique low power (although elect.) display enables use of hand while taking advantage of finger tactile sensitivity. Same speech processing as Nucleus cochlear implant. 7 arif. deafened subjects improved tracking scores from 35 wpm w/o aid to 55 wpm with the aid, after 70 hrs.
SRA-10	Scott, 1980; De Filippo, 1984	Vib. - 3 - chest	Combines info. from 3 freq. bands (hp > 8kHz, 1st format, and 2.4 kHz bp), with unique AM-FM stimulus scheme.	Provides spatio-temporal info about speech signal in a unique display. Arif. deafened subjects - best subj. achieved tracking score of 76 wpm after 10 hrs training (5 hrs aided, + 5 hrs unaided).

Table 3b Summary of Tactile Aids Reviewed (Continued)

types of tactile aids. Thus a system that is good for identifying vowels may not necessarily be useful in understanding connected discourse. Even the widely used tracking test method is subject to variations because it is an interactive test and depends heavily on the cues given by the reader. Thus the aids are difficult to compare directly, but it is possible to get a feel for the types of research that have been done and the range of performance available with current tactile aids.

The generally good performance of the vocoder type tactile aids in isolated word recognition (particularly the work of Brooks and Frost), indicates the utility speech spectral content information. Tactile vocoders have a number of design variables, with respect to the filter bank, that may be varied. For example: what is the frequency range of analysis and the number of filters within this range? Breeuwer and Plomp (1986), reported that auditorily presented information about the speech spectral energy in only two frequency bands, centred at 500 and 3150 Hz, increased lipreading scores dramatically. In Franklin and Saunders' (1984) paper, they presented results that demonstrated the superiority of a 16 channel vocoder device over a 32 channel system. They postulated that the difference was due to the 16 channel device's filters which more accurately modelled the human auditory system. Thus the number of filters may not be as important as the centre frequency, bandwidth and/or temporal response of the filters. The temporal aspects of the

speech signal are stressed in the SRA-10 tactile aid designed by Scott (1974). This device has shown very good results as an aid to speechreading (De Filippo, 1984).

In summary, the performance comparison of various tactile aids is very difficult due to the testing variations. The task chosen to evaluate the benefits of the tactile aid will affect the final result. However, a general conclusion from the body of tactile aid research suggests that the perception of connected speech (sentences) can best be supplemented by timing information, but that the discrimination of words and phonemes in isolation requires spectral information. Thus Scott's approach combining temporal and spectral aspects of speech in the tactile display seems to be most appropriate.

The lack of widespread use of tactile aids by the hearing-impaired population despite the long list of research projects in this area is evidence of the difficulty in designing an effective wearable tactile communication aid. Part of the problem stems from the limited knowledge of speech perception for hearing individuals. Another difficulty and perhaps the greatest challenge, has been not which speech features are best to extract, from the auditory speech perception point-of-view, but which speech features are best suited to transmission via the tactile sense. The capacity of the tactile sense for use as an information channel is often the limiting factor in many of the attempts in providing speech information tactually. Although recent tactile

communication aid research is very promising, it is clear that the limitations of the tactile channel must be carefully examined in order to determine the optimum speech processing and coding strategies.

2.2 Characteristics of the Tactile Sensory Channel

In many communication systems, the information carrying capacity of the channel is often the limiting factor of the system performance and tactile communication aids are no exception. Thus it is desirable to fully characterize the transmission channel in order to design the optimum coding method for the information to be transmitted. The difficulty is that, in the case of tactile aids, the channel is a human sensory process and therefore is subject to variability and difficulty in characterization.

There is a large body of literature documenting various aspects of the tactile sense including tactile physiology, theories of perception and psychophysics. The subsequent discussion of the tactile channel will concentrate on research done with the intent of employing the tactile sense as an information channel. (The reader is referred to the Bibliography for additional information.) In particular, the two common methods for stimulating the tactile sense (or driving the channel), electrical and mechanical, are outlined in Section 2.2.1. Following this, Section 2.2.2 discusses the various parameters available for information coding to a

single tactile excitation source. Section 2.2.3 deals with the effects of multiple transducer coding.

2.2.1 Electrical and Mechanical Stimulation

Two modes of tactile stimulation are available to the designer of tactile aids: electrical and mechanical.

Electrical methods excite the sense of touch by applying electrical current directly to the skin surface. This is accomplished using monophasic or biphasic pulses sent through silver, gold or stainless steel electrodes (Blamey and Clark, 1987; Szeto and Saunders, 1982). Intensity of the sensation is controlled by the width of the pulses.

There are a number of problems associated with direct electrical stimulation, the most important of which is discomfort. Because electrical current is applied directly to the skin, many patients often feel painful discomfort with prolonged use. This was found to be a problem in the tactile speech reception aid developed by De Filippo and Scott and eventually the electrodes were replaced by vibrators (De Filippo, 1984). In contrast to this case, Saunders has been successful in employing electrotactile stimulation in sensory aids for the deaf, without his subjects incurring discomfort (Saunders, 1977; Saunders et al., 1976; and Saunders, 1974). It is interesting to note that despite the fact that Saunders designed the electrocutaneous device for the Scott and De Filippo system, the child subjects found the electrotactile display to be uncomfortable. In addition there is one

commercial tactile aid available, the Tacticon, which uses a 16 channel electrotactile array worn on the abdomen (Franklin and Saunders, 1984).

The magnitude of tactile perception (R) can be expressed in a transduction function, as a factor of the stimulus intensity (I) by $R = kI^\gamma$, where k = a constant and $\gamma \approx 3.5$ for electrotactile stimulation (Matlin, 1988). The discomfort of electrical stimulation of the skin is exacerbated by the very high gamma (for comparison $\gamma \approx .33$ for visual brightness) in the electrotactile transduction function. The result is a limited dynamic range between the threshold of the touch sensation and a painful burning sensation which can occur if the stimulus is too intense (Proctor and Goldstein, 1985). In addition, this transition from a comfortable stimulus to pain can be very rapid (Szeto and Saunders, 1982). Discomfort may also occur if the electrodes do not make good contact with the skin surface and therefore some type of electrode jelly is often required. The physical contact required may not be reliably maintained in larger electrode arrays or for active users, such as children.

Additional problems cited by De Filippo's (1984) tests include: variability in optimum positioning of the electrodes over test trials; the tendency for small changes in electrode location to produce large changes in sensation level; habituation (see next section for details) and the problem of parental approval for electrical stimulation of their deaf children. A further difficulty with electrotactile

stimulation stems from the high voltages required, on the order of 100+ V, to provide the necessary current. Special precautions must be taken to isolate and current limit the source in laboratory versions and portable versions require voltage conversion from battery voltage levels.

Despite all of the difficulties associated with direct electrical stimulation, these systems are of interest to designers of tactile sensory aids due to the low power requirements. Electrotactile stimulators typically require 2-10 times less energy than vibrotactile transducers and thus are well suited to wearable prosthetic systems. Also, the small size of the electrodes, particularly in the thickness dimension, allow for novel placements of electrode arrays (Blamey and Clark, 1987).

The disadvantages of vibrating type transducers, namely size and power consumption, has for the most part, limited the use of vibrotactile stimulation to the laboratory. Within the last decade, however, progress has been made in the implementation of miniature, low power vibrating devices (Audiological Eng. Corp., 1986, Sherrick, 1984; Ifukube, 1982). There are two types of vibrating transducers that are commonly used in tactile aids.

The first is the electromagnetic type which consists of a miniature coil of wire and a metal core and/or surround to complete the magnetic circuit. When electrical current is passed thorough the coil, either the core moves (United States Patent #4,728,934, 1988) or the coil itself moves

between metal endplates (Audiological Engineering V1220 Vibrator). The vibration is transferred to the skin by a contact point connected to the moving part, or simply by the shaking of the case of the transducer. The electromagnetic devices are the most popular in the commercial single channel tactile aids because of their durability and ease of use. The power required to produce sufficient vibration is often relatively large, thus limiting the number of transducers that may be employed in a wearable system.

The second type of vibrating device uses piezoelectric materials in a multilayer, reed-type configuration. When a biphasic signal is applied to the reed, it bends back and forth, as the polarity of the signal changes. A contactor pin is usually attached to the end of the reed and pushes against the skin in order to produce the sensation (Ifukube, 1982; Sherrick, 1975). The piezoelectric vibrators require very little power, but like the electrical stimulators, they do require very high voltages of up to 200 V. The other disadvantages of the reed benders is that wearable versions tend to be fragile and larger in size than the miniature electromagnetic devices.

In terms of comfort, ease of use, durability and expediency, the commercially available, electromagnetic vibrators, such as the V1220 model from Audiological Engineering Corporation, seem to be the best choice (Frost, 1987). Despite the inherent difficulties of direct electrical stimulation, a few investigators have had some success with

electrocutaneous tactile channel stimulators. In particular, the novel arrangement of Blamey's electrotactile array, combined with thorough testing of such systems (Blamey and Clark, 1987) has demonstrated the feasibility of direct electrical stimulation. However, the associated problems, such as day-to-day variations, electrode coupling and overall acceptance still remain. The piezoelectric devices combine the low power requirements of direct electrical stimulation, with the advantages of electromagnetic transducers. With the exception of devices that stimulate the fingertips, piezo transducers are larger than other similar transducers. In addition the piezoelectric vibrators tend to be fragile and thus are not suited for tactile aids to be worn by young children. Whichever method is chosen to drive the tactile channel, the characteristics of the channel must be investigated in order to optimize the coding of information to be sent through the channel.

2.2.2 Single Source Tactile Channel Characteristics

For a single excitation source, there are a number of tactile channel characteristics which are important for information transfer. The basic stimulus parameters available to the designer of tactile sensory aids, are: location, intensity, frequency and duration (Geldard, 1960). The first consideration is the location of the transducer on the body. The selection of stimulator location is usually based on the sensitivity to vibration. In addition, for tactile sensory

aids such factors as comfort, convenience, ease of use and aesthetics, all must be considered.

Table 3 lists the sensitivities of various body sites to three parameters. The column of vibratory thresholds ranks the finger tip to be the most sensitive. This is reinforced by the success of many blind people at reading braille characters. The obvious disadvantage of using the fingertips in a tactile aid, is that one or both hands must be dedicated to the reception of speech. Blamey and Clark (1987) have avoided this problem by electrically stimulating the nerves that run from the finger tips, along the sides of the fingers, thereby obtaining sensitivities nearing those on the fingertip while allowing near normal function of the hand. The sternum and forearm also exhibit low vibratory thresholds as well as allowing the transducers to be worn inconspicuously. As a result, a number of tactile aids have used these locations.

Sensitivity to vibration intensity is a very widely used dimension for the coding of information in tactile sensory aids. In addition to the absolute threshold for the sense of touch, it is of interest to note how this percept varies with frequency. Verillo and his colleagues have done extensive work in the area of tactile perception and have determined what are known as the equal sensation contours for the tactile channel (Verillo, 1970).

Two-point limen (mm)		Point-Localization error (mm)		Vibration (dB re 1µm @ 200 Hz)	
Mid-finger	2.5	Index finger	1.4	Palmer finger	-23.1
Index finger	3.0	Nose	1.5	Palm	-23.1
Thumb	3.5	Ring finger	1.6	Hand dorsum	-19.1
Ring finger	4.0	Mid-finger	1.7	Sternum	-11.1
Little finger	4.5	Thumb	1.8	Volar forearm	-8.2
Upper lip	5.5	Upper lip	1.9	Sole	-6.9
Cheek	7.0	Little finger	2.0	Hallux	-2.3
Nose	8.0	Hallux	2.1	Post, upper arm	0.8
Palm	11.5	Cheek	2.2	Elbow	2.3
Hallux	12.0	Forehead	7.0	Medial ankle	2.9
Forehead	15.0	Palm	7.1	Thigh	5.1
Sole	22.5	Sole	7.5	Lips	6.8
Belly	34.0	Belly	9.0	Lower leg	8.0
Breast	36.0	Forearm	9.4	Out upper arm	10.1
Forearm	38.5	Breast	9.5	Lower abdomen	13.1
Shoulder	41.1	Shoulder	9.6	Nose	13.4
Back	44.0	Upper arm	10.5	Chin	15.0
Upper arm	44.5	Calf	11.0	Kneecap	15.0
Thigh	45.5	Thigh	11.5	Calf	15.0
Calf	47.0	Back	12.5	Buttocks	22.9

Table 4 Sensitivity of Body Sites (after Kenshalo, 1978)

Figure 7 shows the equal sensation contours for 10 frequencies and 7 sensation levels. As illustrated by the dip in the curves, subjects were most sensitive to the vibrations when the frequency was 250 Hz and a much stronger signal was required to create the same sensation level at lower frequencies. This frequency dependant intensity sensitivity has been found by other researchers for both vibrotactile (Bernstein et al., 1986; Verillo, 1985) and electrotactile stimulation (Anderson and Munson, 1951). Thus most tactile sensory aids drive the transducers with a signal frequency of 150 to 400 Hz.

Figure 7 Equal Sensation Curves (Verillo, 1970)

The static force applied to a vibrator has been shown to affect the intensity sensation with an increase in static force causing a decrease in the lowest perceptible intensity (Lamore and Keemink, 1988). Bernstein et al. (1986), reported that the vibrotactile sensitivity thresholds for hearing children, two prelingually, profoundly deaf children and adults differed very little across subjects, for frequencies above 50 Hz.

When sensation intensity is used as an information coding dimension, the maximum comfortable level and thus the dynamic range of the tactile channel is of interest. Also the number of discriminable amplitude steps within this amplitude

range is also important. In the case of electrotactile transducers, Blamey and Clark (1987), report dynamic ranges of between 12 and 18 dB, depending on the subject. The size of the smallest discriminable step, or difference limen, was found to be approximately 0.36 ± 0.12 dB. Other studies of electrical stimulation have produced similar results (Anderson and Munson, 1951; Hawkes, 1961; Hawkes and Warm, 1960; Brown et al., 1966; Saunders, 1974). Note that the dynamic range available through electrical stimulation is approximately 1/8 th that of the auditory system. Thus some type of speech coding will be required just to match the dynamic range of the auditory system and speech signal to the tactile information channel.

In addition to the intensity dimension, the frequency of vibration may also be employed to code speech information. As was previously illustrated in Figure 7, the optimum frequency of vibration, in terms of amplitude sensitivity is approximately 250 Hz. For frequencies above and below this, the tactile sensitivity to intensity decreases. Thus, a change in frequency may be perceived only as a change in amplitude of the vibration. Conversely, it has been reported that increases in the amplitude of vibration sometimes create the sensation that the frequency of vibration is decreasing (von Békésy, 1957; 1959). Despite these difficulties in employing vibrotactile frequency in the encoding of speech information, the interest remains due to the importance of frequency in the speech signal.

Rothenberg et al. (1977), examined the range of vibrotactile frequencies and corresponding perceptions of amplitude for a number of different waveforms. Their results for sinusoidal signals are similar to those of Verillo, shown in Figure 7. However a signal consisting of a train of Gaussian shaped pulses provided equal intensity sensations at all frequencies over the range of 20 - 500 Hz, when the amplitude of the vibrations was 14 dB SL above the threshold level (@ 250 Hz). Their experiments on the difference limen, or minimum detectable step, for frequency suggest that there are at least 7 detectable steps when vibrotactile frequency is increased from 20 to 500 Hz. Rothenberg et al. (1977) postulated that, with the appropriate transformation, it would be possible to effectively code fundamental frequency as vibrotactile frequency.

The difficulty however, is that test of vibrotactile frequency perception with prelingually deaf subjects yielded very poor results, presumably due to their unfamiliarity with auditory frequency discrimination (Rothenberg et al., 1977). Indeed, later tests done with a vibrotactile pitch display, employing pitch to frequency of vibration coding, showed very poor results when tested with deaf subjects (Rothenberg and Moliter, 1979). Thus, vibrotactile frequency is not a widely used dimension in the coding of information for the tactile channel.

Closely related to the frequency resolution characteristics is the temporal processing ability of the tactile sense.

There are a number of ways to classify the temporal capacity of the tactile channel, but two particularly useful schemes in terms of the tactual transmission of speech are the temporal difference limen and the temporal modulation transfer function.

Temporal difference limen is a measure of the minimum detectable time difference between two stimuli. Gescheider (1973), reported threshold on the order of 8-10 ms using an electromagnetic vibrator. Geldard and Sherrick (1971), used piezoelectric bimorph stimulators with much faster impulse responses than Gescheider's device and found thresholds of between 3 and 5 ms. Gescheider also reported that there was an interaction between the temporal difference limen and the intensity of the vibration and the minimum detectable time difference decreased with increasing vibration amplitude.

Another means of estimating the temporal transmission capacity of the skin is the temporal modulation transfer function. This transfer function is used to describe a system's response to a time-varying input. Typically the test signal consists of a carrier waveform with the amplitude modulated over time. Weisenberger (1986) found that for a 250 Hz sinusoidal carrier signal the tactile sense was most sensitive to modulation at a frequency of 40 Hz but subjects could detect modulations down to ~10 Hz range and up to ~80 Hz. This type of information is valuable to the design of tactile aids for the deaf because a number of vibrotactile coding strategies employ amplitude modulation. Weisenberger

concluded from her data that the transmission of speech at a phoneme rate is feasible, but information about rapid temporal variations in speech, such as stop consonants, may not be resolved by the tactile channel.

Given the limitations of the tactile channel in terms of amplitude, frequency and temporal resolution, the transmission of speech information will require extensive processing and coding in order to suit the channel characteristics. The simple single channel tactile aids for the hearing impaired can only transmit limited speech cues to the user due to the channel limitations. In order to increase the amount of speech information imparted to the user, many systems use more than one vibrator or electrode.

2.2.3 Multi-Source Tactile Channel Characteristics

Each stimulator in multi-source tactile display schemes is subject to the above limitations. However, forming arrays of transducers creates another dimension for coding information; that is, spatial location. There are problems associated with the use of multiple output devices and these problems involve interactions among the various transducers. This is known as vibrotactile masking.

The first difficulty relates to the identification of the active vibrator. In Table 3, the columns labelled point-localization error and two-point limen contain information pertaining to the task of identifying the location of stimulation. Point localization tests a subject's accuracy in

pointing to the location of a static touch. As shown in the Table, subjects can identify the location of vibration to within 1.4 mm on the fingertip in the best case. At the other common site of tactile stimulation, the forearm, the location of the touch could only be identified to within 9.4 mm.

When two stimuli are presented simultaneously there exists a minimum spatial distance between the two at which two stimuli can be identified. This is called the two-point limen or threshold. For spacings less than the two-point limen subjects perceive only one sensation. For example, from Table 3, the two-point limen on the forearm is 38.5 mm and therefore, two identical static forces spaced less than 38.5 mm apart, will combine to form one sensation, subjectively located midway between the two forces. Although the figures listed in Table 3 pertain to static forces applied to the skin, the two-point limen tests also apply to vibrations. However, the interaction between vibrators is more complex.

In the case of vibratory or electrical stimulation the combination of two transducers spaced less than the two-point limen apart forms what is known as a phantom sensation (Tachi et al., 1980; Alles, 1970; McEntire, 1972). By varying the relationship of stimulus amplitudes and onset times of the two vibrators, the perceived sensation can be positioned at different locations between the two transducers (Alles, 1970). This may also be extended to two dimensions by using three transducers arranged in a triangle (Tachi et al., 1980; McEntire, 1972).

Closely related to the phantom sensation, is the phi phenomenon, where the sensation is perceived as moving or "gouging" between two transducers (Kenshalo, 1978). The rate of movement depends on the transducer spacing and the interstimulus onset intervals. In addition to the continuous movement sensation, a hopping of discrete points of sensation can be created by multitransducer systems. This feeling "as if a tiny rabbit is hopping along the skin" (Geldard and Sherrick, 1972, p.178), is created with three vibrators in a linear array and square pulse signals with constant temporal spacing between them (Kenshalo, 1978; Craig and Sherrick, 1982). There has been some effort in employing these two schemes in the coding of information, but the variations among different subjects to the phantom sensations and the difficulty in generating them reliably has limited their use in tactile aids (McEntire, 1972; Tachi et al., 1985).

As well as producing phantom sensations, interactions among the transducers in an array may also tend to obscure some sensations. This type of interaction is termed masking and occurs when one stimulus interferes with the detection of another stimulus. For example, when more than two transducers are vibrating, not all of them will be detected due to the complex relationship among them which produces masking (Craig and Sherrick, 1982; Craig, 1985). Spacing the stimulators farther apart and increasing the time between adjacent stimulations reduces masking. However, it also confounds the two goals of tactile information displays: i)

high information rate and ii) compact displays. Some researchers suggest that masking may not necessarily be detrimental in displays of speech since it is an indication of the interaction among stimuli. The interaction may produce new spatiotemporal patterns, through mechanisms of integration, resulting in better transmission of information (Kirman, 1973).

Tactile display systems which employ a large number of transducers have a number of more practical problems as well. First of all, the power consumption of such systems tends to make them impractical in wearable systems, particularly if the transducers are of the electromagnetic variety. In addition, the physical size of such displays may make them conspicuous and uncomfortable to wear. Due to the complex interactions among stimulators these types of displays may be difficult to learn to use. Finally, the interactions among transducers may lead to difficulties in evaluating the performance of such tactile sensory aids since the experimenter may be unaware of the effects of the interactions on the subjects perception of various features.

2.2.4 Tactile Channel Capacity: Conclusions

As was the case with the comparison of tactile sensory aids, one must exercise caution in comparing the data of various research groups in the area of tactile channel characteristics. Differences in experimental setup and procedures, subject differences and varying amounts of

training can make direct comparisons of results invalid.

The most important question is whether or not the capacity of the tactile channel is sufficient for the transmission of speech information to the profoundly deaf. As outlined in the preceding section, there are a number of limitations associated with the tactile channel but the success of some of the tactile aids reviewed in Section 2.1 suggests that with proper coding of the speech signal the tactile sense can be used to convey at least one feature of the speech signal successfully.

Further, the speech comprehension achieved by deaf-blind persons using the "Tadoma" method of speech reception, where the 'listener' places his/her hand on the speaker's face and throat, provides the evidence that speech may indeed be transmitted via tactually (Reed et al., 1985; Norton et al., 1977). Thus there is evidence that speech communication can occur via the tactile channel, although electronic tactile aids have yet to achieve the performance that some people have reached using the Tadoma method.

2.3 Speech Processing Strategy

Information from Sections 2.1 and 2.2 dealing with tactile sensory aids and tactile channel characteristics respectively, can be used to determine an optimal speech signal processing strategy. In order to be successful, the speech processing method should be based on both the speech characteristics required to supplement communication and the

capacity of the tactile channel, through which the information will be transmitted. The preceding overview clearly indicates that the implementation of any speech processing strategy is subject to many variables. Thus an implementation platform should allow modifications to the algorithm and customization to the patient in an efficient manner.

The speech amplitude envelope extraction systems, such as the Tactaid I and Minivib tactile aids, demonstrated the significance of timing information as a supplement to the speechreading task. The Queen's University, 16 channel tactile vocoder allowed subjects to perform very well in the word recognition tasks without lipreading, thereby verifying the important role of spectral information for interpreting speech. Breeuwer and Plomp (1986; 1984) reported that the low-high frequency content indication performed better than formant frequency information in supplementing speechreading, in their experiments. The indication of the dominate energy frequencies in a broader sense than the formant peaks provided subjects with a better assistive cue to lipreading than the finer information about individual formant peaks. Thus, information about the spectral tilt (low-high frequency content indication) of the speech signal is an important supplement to lipreading.

A speech processing strategy should therefore, attempt to impart both timing information and spectral content. The success of Scott's tactile aid (De Filippo, 1984) which

provides this spectral/temporal combination is a further indication that both time and frequency speech information are important.

The unsuccessful tactile aids may also make a contribution to the design of the present system. In particular, the difficulties encountered with the two-dimensional tactile output displays, suggests that this is not the approach to take. These systems tend to be very difficult to use and rather cumbersome to wear. Limits in wearability are due to both the physical size of the display, as well as the power consumption.

The performance of the various tactile aids combined with the research in the area of tactile channel capacity, provides an indication of the limitations and requirements for efficient information transmission via the tactile channel. Many of the more capable tactile aids all share a similar stimulus signal consisting of a amplitude modulated sinusoidal or square carrier waveform with a frequency in the range 100 - 400 Hz. The tactile channel research supports this choice of signal due to the maximum tactile amplitude sensitivity at approximately 250 Hz.

The limited information transmission capacity of the tactile channel requires the use of multiple tactile transducers, in order to use the spatial location as an information coding parameter. But due to the problems of interactions among transducers, power consumption and size, the number of transducers should be kept to a minimum.

Tactile perception research has demonstrated that intersubject differences can often be large, particularly with electrotactile systems (Blamey and Clark, 1985). Long term training also tends to alter subjects' sensitivity to various tactile parameters. Therefore a programmable system for tactile stimulation will be necessary. Programmability also allows a system to take advantage of the latest tactile perception research, without major redesigns of electronics.

The principal component analysis of speech spectra (PCA) meets many of the above design requirements, both in terms of the necessary speech information and the limitations of the tactile channel. Chapter 3 is devoted to the PCA technique and its application in a tactile aid for the profoundly deaf. Briefly, PCA is applied to the outputs of a bandpass filter bank whose properties are based on human audition. The effect of this analysis is to transform the large number of filter outputs (typically 16) into new dimensions called principal components (PCs). It has been found that as few as four principal components are an accurate representation of the original speech signal (Zahorian and Rothenberg, 1981). In terms of previous speech processing schemes, the first PC corresponds to the amplitude envelope of the speech signal and the second provides a low-high frequency indication. In addition, as few as four transducers are required to drive the tactile channel and the effective amplitude modulation of a 250 Hz carrier signal scheme may be used. Thus principal components analysis seems well suited to a tactile aid.

Chapter 3

3.0 Introduction to Principal Components Analysis

As outlined briefly in the previous chapter, the principal components analysis of speech spectra is well suited to use in a tactile communication aid. The principal components include many of the more successful speech features used in previous systems as well as meeting the requirements for efficient transmission via the tactile channel. The particular application of principal components analysis to speech processing has been examined by a number of researchers and is presented in Section 3.2. The configuration and variables involved with the use of PCA of speech for a tactile aid is discussed in Section 3.3. First, a general description of principal components analysis is provided.

3.1 Principal Components Analysis (PCA)

Principal components analysis is a statistical technique developed by Hotelling (1933), as a means to find an efficient representation of a set of correlated data. In scientific experiments and studies that produce a large number of correlated variables, the PCA algorithm has been used as a method of data reduction. If for example, an investigator is interested in the deviations of the original data, it would be advantageous to reorganize the data into new variables that are ordered according to the total amount

of variance accounted for. The number of variables or dimensions required for analysis may then be reduced since a large percentage of the variance of the original data will now be accounted for by a smaller number of the new variables. The statistical theory and background of principal components analysis has been well documented (Hotelling, 1933; Anderson, 1963; 1971; Papoulis, 1965; Ahmed and Rao, 1975; Harman, 1976).

In the communication theory literature, the principal components analysis is often referred to as the Karhunen-Loeve expansion or transform (KLT). The KLT derives an optimal set of orthonormal basis vectors to represent a set of zero-mean data (Allerhand, 1987). The new basis vectors, (or principal components) are optimal in the sense that for a N -dimensional signal space, the difference between the original signal and one reconstructed from the first P basis vectors, where $P < N$, is minimal according to a mean-square error criteria (Bustamante, 1986). Mathematically, the PCA and KLT are identical (Chan and Cheung, 1986). The principal components analysis (PCA) nomenclature will be used in the remainder of the thesis.

Principal components analysis then, attempts to reorganize multidimensional data in a manner that eliminates correlations among the various dimensions. Searle (1985) and others have described this process as a geometric rotation of the original N dimensions to form new dimensions, the first

few of which account for a large percentage of the variance of the original data.

A simple example is shown in Figure 8 for data having three dimensions. The original dimensional axes are x , y and z , as shown by the dashed lines. For the data plotted, a more optimum set of axes may be found by rotating x, y, z to position x', y', z' . As depicted in Figure 8, the data may now be represented by a single dimension, defined by the x' axis.

Figure 8 Three Dimensional Rotation Example

In the case of speech processing, consider the outputs of a 16 channel filter bank, with speech input. Each spectral frame may be thought of as a point in 16-dimensional space.

An utterance then, is a trajectory through this space. Based on this utterance, a new representation may be found by rotating the axes until the maximum variance appears along one axis. This is the first new coordinate. Next the data is rotated again, in order to find a second dimension, orthogonal to the first, which accounts for the next largest variance. This process is repeated for all remaining dimensions. The direction cosines of the new coordinate system, with respect to the old, are the principal component basis vectors.

Mathematically, the basis vectors for the utterance are the eigenvectors of a covariance matrix formed with the spectral points of the utterance. The eigenvector with the largest eigenvalue corresponds to the first basis vector and accounts for the most variance of the original signal. The eigenvector with the next largest eigenvalue, accounts for the second most variance and so on.

In the case of a filter bank, with N channels, a vector of spectral data, at time t , may be formed as:

$$x_t = (x_1, x_2, \dots, x_N); \quad (1)$$

where: x_n is the amplitude of spectral energy in bandpass filter n .

The covariance matrix C , of the spectral data frames over K frames is (Ahmed and Rao, 1975):

$$C_{ij} = 1/K * [\sum_{k=1}^K (x_{ki} - \overline{x_i})(x_{kj} - \overline{x_j}), \quad (2)$$

for $i, j = 1, 2, \dots, N;$

where: x_{ki} = i^{th} data sample for the k^{th} frame
 $\bar{x}_i$ = average of i^{th} data sample over k frames
 N = number of data elements in each frame
 i.e. the number of filters in the bank
 K = total number of data frames.

The eigenvectors Y_{ij} , are calculated from the symmetric $N \times N$ covariance matrix C_{ij} . The principal component basis vectors are the N eigenvectors, of length N : $Y_{i1} \dots Y_{iN}$. The principal component representation, PC_i , of the spectral data in a frame at time t is formed by calculating a linear combination of the original frame data, $x_1 \dots x_N$, using the elements of the basis vector as weights for each spectral channel (1 through N), as shown in equation (3):

$$PC_i = Y_{i1}x_1 + Y_{i2}x_2 + \dots + Y_{iN}x_N; \quad \text{for } i = 1 \dots N \quad (3).$$

Thus, rather than the N dimensions required to represent the spectral data at time t , P principal components may be used, where $P < N$. In many problems to which principal components analysis is applied, the number of principal component dimensions required to represent the original data is 25% or less than the number of dimensions in the initial data. This increases coding efficiency.

The difficulty in implementations of this algorithm is that the time required for covariance matrix formation, eigenvector calculation and linear combination of the original data precludes the application to many situations that require real-time solutions. Fortunately, the use of

principal components analysis in speech spectra coding produces basis vectors that stabilize and remain constant after 0.5 - 10 minutes of speech has been processed (Zahorian, 1988a). Thus the principal components basis vectors need to be calculated only once, prior to real-time operation. In fact, as will be discussed in the following section, the basis vectors for speech have been shown to be invariant to speaker sex, age and even language in some cases.

3.2 Principal Components Analysis of Speech

The application of PCA to speech processing has been investigated by a number of research groups, beginning with Kramer and Mathews (1956). The real utility of PCA in real-time speech processing arises from the similarity in the basis vectors derived by different research groups. Recall that the elements of the principal components basis vectors are used to weight the spectral channels in forming the linear combinations that are the PC representation of speech. Thus if the channel weights necessary to calculate each PC are constant, it may be possible to perform the conversion to PCs in real-time for speech spectral data.

In studies done by a group at the Institute for Perception TNO (Plomp, Pols and van de Geer, 1966; Klein, Plomp and Pols, 1970) and by other groups (Li et al., 1969; Taylor and Pols (in Searle, 1985); Zahorian, 1981; 1988; Lennig, 1984; and Searle, 1985), the basis vectors derived

from speech spectra were very similar, both across the studies and for different speakers within each research group. These principal component basis vectors approximate a $1/2$ cosine series. Figures 9 through 13 show the basis vectors found experimentally by Searle's group (a) (1985), Taylor and Pols (b) (in Searle, 1985), Zahorian (c) (1981) and the $1/2$ cosine series (Cos Series).

Figure 9 Principal Component Basis Vector #0

(a) Searle (1985)

(b) Taylor and Pols (in Searle, 1985)

2

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS
STANDARD REFERENCE MATERIAL 1010a
(ANSI and ISO TEST CHART No. 2)

Figure 10 Principal Component Basis Vector #1

- (a) Searle (1985)
- (b) Taylor and Pals (in Searle, 1985)
- (c) Zahorian (1981)

Figure 11 Principal Component Basis Vector #2

- (a) Searle (1985)
- (b) Taylor and Pals (in Searle, 1985)
- (c) Zahorian (1981)

Figure 12 Principal Component Basis Vector #3

(b) Taylor and Pals (in Searle, 1985)

(c) Zahorian (1981)

Figure 13 Principal Component Basis Vector #4

(b) Taylor and Pals (in Searle, 1985)

(c) Zahorian (1981)

The Taylor and Pols data demonstrated the similarity of the basis vectors for male and female speakers and for English and Dutch languages. Note in Figures 9 through 13, the gross shapes are similar, but in some cases the amplitude scaling or sign may be different, or there may be a shift left or right, as compared to the $1/2$ cosine series. Some of the variation between groups is due to differences in filter bank characteristics, such as number of filters, frequency range of analysis and the response of the filters themselves (Zahorian, 1988b).

Taylor and Pols also calculated the variance of the original spectral data accounted for by the PCs. Figure 14 plots the variance of the filter bank outputs and the variance of the principal components. Note that over 90% of the variance is accounted for by the first four PCs. Plomp et al. found that 84.5% of the variance was accounted for by the first four PCs for the Dutch language. One might question the relevance of measuring the speech information retained by the principal components in terms of variance of the speech spectra.

Figure 14 Variance Accounted For By Principal Components (after Searle, 1985)

Perceptual experiments performed by Kramer and Mathews (1956), Klein et al. (1970) and Zahorian and Rothenberg (1981), tested the information bearing content of the PCs by resynthesizing speech from the principal components representation and found good intelligibility when as few as four PCs were used to reconstruct the speech. Zahorian and Rothenberg (1981), reported that 95% of words in unfamiliar sentences were correctly identified, when the speech signal was resynthesized from the first 4 PCs. The performance of the PC speech coding technique was compared to that of Linear Predictive Coders (LPC) having the same number of coded parameters. The best PC vocoders were at least as good and

sometimes better than the LPC systems in terms of the intelligibility of the reconstructed speech (Zahorian and Rothenberg, 1981).

PCA of speech spectra has also been used in speech recognition tasks. Klein et al. (1970) achieved a 98% correct rate for computer recognition of Dutch vowels using only four principal components. Pols (1971) used three PCs in a template matching speech recognizer and obtained a machine recognition score of 98.8% correct, for 20 utterances (the digits 0-9 and 10 computer commands). In the high noise environment of a simulated fighter jet cockpit, Doddington and Naik (1985), found that a Principal Component based speech recognizer exhibited an error rate one half that of a typical LPC speech recognizer. Naik and Doddington (1986), also successfully applied principal components analysis of speech to a speaker recognition task.

The results of the various applications of PCA to speech coding and recognition verify that a small number of the principal components of speech do contain a large amount of the original information carried by the speech signal. In terms of perceptual characteristics, Pols et al. (1969), demonstrated that there is a high correlation between "principal component space" and "perceptual space", with correlation coefficients as high as 0.992, 0.991 and 0.742, for the first three PCs, respectively.

Yilmaz (1967; 1968) suggested a novel theory of speech perception and discussed a number of interesting, although

informal, experiments to verify his hypothesis. Yilmaz based his theories of auditory perception on visual/auditory perception similarities. Despite the obvious differences in the visual and auditory perceptual mechanisms, Yilmaz worked on the similarities, namely the perception of wideband noise characteristics of whispered speech and the visual perception of wideband radiation, seen as colour. In the development of his theory Yilmaz considered two areas:

- a) the physical properties of the stimulus energy and
- b) the evolutionary history and needs of the organism for survival.

From the properties of the speech signal and the evolution of the human auditory system, which he postulates developed to be optimal for speech reception, Yilmaz suggested a two-dimensional representation of speech. This auditory space representation was argued to be analogous to the colour triangle used to mathematically represent the perceptual phenomena of colour mixture of lights. The interesting aspect of the Yilmaz theory is that the basis vectors for the auditory space were developed from theory to be a sine and cosine vector for each of the two dimensions. Thus Yilmaz derived a speech representation similar to that of PCA, based on the physical characteristics of the speech signal and the evolution of the auditory perceptual system. The PCs of speech retain the information content of the signal in an efficient manner, but may also be important to the perception of speech in our auditory system. Traditionally, the speech

variables thought to be important for speech understanding have been the formants and timing information.

It is interesting to compare the relationship of the principal components of speech spectra formed using the basis vectors in Figures 9-13, to traditional speech features. For example, the first PC of speech is the sum of the filter bank energies, all weighted equally, and, hence, corresponds to the amplitude envelope of the speech signal. The second basis vector weights in Figure 10 will indicate the tilt of the speech spectrum; that is, whether the dominant frequency content is high, as in consonant type sounds, or low, as in most vowel sounds. Similarly the other PCs are spectral shape factors giving information about more specific spectral shapes with each additional PC. Klein et al. (1970) found that the second and third PC basis vector sums had very high correlation coefficients, 0.997 and 0.945, with the first two formants, respectively, of speech.

The principal components representation of speech spectra allows a data reduction over standard filter bank analysis while, also, preserving the aspects of speech that have been shown to be important for speech perception. The speech features that were successfully used in tactile communication aids for the deaf are also well represented in the PCs of speech. Given this, the PCA of speech for use in a communication aid for the deaf seems to be a logical choice.

Zahorian (1988; Zahorian and Jagharghi, 1985) have used PCA in a colour speech display as a speech production

training aid for the deaf. In their system, the PCA of the filter bank energies was transformed to the Red, Green, Blue (RGB) parameters of the colour display. The transformation from PC space to the colour display was arranged to optimize the separation of three widely spaced vowels /u/, /i/ and /a/, each of which was assigned one colour. Other vowels appeared as secondary colours, depending on the relationship the three vowels assigned the primary colours. Normal hearing subjects could identify and match the vowels produced on the display with accuracies of 93% and 85%, respectively for the two tasks. Evaluation with deaf subjects is currently being carried out.

The PCA of speech spectra results in a small number of slowly-varying parameters that retain a large amount of the original speech information. The weighting functions (PC basis vectors) used to determine the PCs of speech are invariant across speakers and, apparently, across some languages. In addition, the PCs relate well to the speech features that have been employed successfully in tactile sensory aids for the profoundly deaf. The implementation of PCA of speech for use in a tactile aid would seem to be a worthwhile venture.

3.3 Configuration of PCA for a Tactile Aid

There are a number of variables in the PCA algorithm that must be addressed, prior to implementation of the algorithm in a tactile aid. The important variables relate to

the spectral analysis of the speech and the weighting functions used to determine the PCs. In the spectral analysis procedure variables include: speech pre-emphasis, centre frequency and width of the frequency bands in the analysis and frequency range used. In the formation of the PCs from linear combinations of the spectral energies, the use of non-linear energy scaling and the actual weighting functions must be decided upon.

3.3.1 Spectral Analysis Variables

Most of the PCA systems discussed in the previous section used spectral analysis techniques with centre frequency and bandwidth characteristics that approximate the spectral processing in the human auditory system. The main characteristics are logarithmically spaced centre frequencies (f_c), coupled with bandwidths proportional to the band f_c . Thus the bandwidths of analysis increase with frequency. The advantage of such a scheme, when implemented as filters, is that the high frequency channels have very rapid temporal response, due to the wide bandwidth. This is necessary to track the rapid spectra changes of consonant sounds. The lower frequency bands have longer response times, but narrower bandwidths which allows the tracking of vowel sound formants.

The Queen's University tactile vocoder enabled users to achieve very good word recognition scores without speech-reading, using a filter bank with the above characteristics.

In fact the 1/3 octave bandwidth filter bank used at Queen's was originally developed by Searle et al. (1979) and used successfully in a stop consonant recognition system. Searle and his colleagues, based their filter designs on the auditory nerve fiber experiments done by Kiang and Moxon (1974). To a first order approximation, the peripheral auditory system has a frequency resolution of 1/3 octave above 400 Hz. Searle has found that a satisfactory implementation of the filters is in the form of 6th order Butterworth type models (Searle, 1985).

In their work on stop consonant discrimination, Searle et al. (1979) found that the time constants of the energy detectors that followed each bandpass filter had to be inversely proportional to the BPF bandwidth. In this manner, the wide bandwidth, high frequency, fast impulse response, bandpass filters have small time constant detectors in order to resolve the rapidly changing high frequency features present in the consonant portions of speech. The fast energy detectors also agree with Kiang's data on the response of individual auditory nerve fibers to acoustic clicks. Thus, the spectral analysis procedure must have fast temporal response in the high frequency bands in order to resolve consonant information in speech signals.

The filter bank used by the group at Queen's University spans a frequency range of 160 to 6300 Hz. The high frequency band will require a speech sampling rate of approximately 16 kHz if the system is to be implemented digitally. At the 16

kHz sampling frequency a real-time implementation may not be feasible, at present. Thus a digital PCA system may have to limit the upper frequency of analysis. This may not be critical since Breeuwer and Plomp's speechreading experiments (1984), showed that information in the 3.15 kHz frequency band (with no other higher frequency bands present), was an effective supplement to lipreading. In addition, the perception of telephone-quality speech is not overly impaired by the approximately 3.4 kHz upper frequency limit.

Prior to the spectral analysis unit some research groups include a 6 dB/octave pre-emphasis function above 500 Hz. It has been found that there is an overall -6 dB/octave rolloff with increasing frequency for speech radiated from the lips (Witten, 1982). The pre-emphasis function flattens the overall speech spectrum prior to the spectral processing. Scaling of spectral energies after spectral analysis is discussed in the next section.

3.3.2 Principal Component Analysis Variables

Apart from the actual weighting functions, the other variable in the PCA of speech is the amplitude scaling of the spectral band energies. Zahorian and his colleagues (Zahorian, 1988a; Zahorian and Rothenberg, 1981) have experimented with a number of linear and non-linear scalings of the spectral energies prior to calculating the PCs. A logarithmic scaling was found to perform the best. This

logarithmic weighting also approximates the perceptual amplitude scale of the auditory system.

The last variable in the PCA implementation is the actual basis vector values and the number of PCs retained for transmission. As shown in Figures 9-13, the first five basis vectors, found by different research groups are similar to one another. There are, however, slight differences which may be due to variations in filter bank configurations (centre frequencies and bandwidths), pre-emphasis and the duration of speech analyzed. Note, however, that the basis vectors in Figures 9-13 approximate a $1/2$ cosine series. Searle's group (Searle, 1985) have successfully used the $1/2$ cosine weights in much of their work and Zahorian has adopted this configuration as well (Zahorian, 1988a). The $1/2$ cosine series summations will, therefore, be used to calculate the PCs of speech in the prototype tactile implementation.

The number of PCs to be calculated is limited by the tactile channel and corresponding transducers used. As outlined in Chapter 2, the optimal tactile display will use a small number of transducers in order to reduce power consumption and simplify the interpretation of the tactile signals. Experiments in speech coding and recognition have shown that the first four PCs retain a large amount of the original speech information. Therefore, the first four principal components of the speech spectra will be calculated, for transmission via the tactile channel.

3.5 Summary of PCA for Use in a Tactile Aid

PCA is a promising speech processing algorithm for use in a tactile aid, both in the speech features represented by the principal components of speech spectra and by the low number of parameters necessary to transmit the information via the tactile channel. The speech features that performed best in previous aids for the profoundly deaf, namely envelope amplitude, low/high frequency distinction and spectral energies, are well represented by only four PCs. Thus the number of tactile transducers required is low and the resulting tactile patterns will be relatively simple.

The algorithm itself, requires the transformation of the speech signal to the spectral domain and four linear combinations of the spectral energies. The weights for the linear combinations and the characteristics of the spectral transformation are fairly well defined in the literature. The signal processing requirements, although intensive, do allow a compact real-time system to be implemented. The details of this follow in Chapter 4.

Chapter 4

4.1 System Level Outline

Following a current system design standard, the development of a tactile aid has involved simulations of the speech processing algorithms on a floating-point computer (PC, Mini, Mainframe, etc.), usually not in real-time. This thesis departs from this standard and proceeds with immediate implementation of a compact, fixed-point, real-time system. There are two very important reasons for bypassing the simulation step. First, the performance and characteristics of the PCA algorithm, as it applies to speech, are well documented. Given the results of past research in this area and the suitability of PCA in terms of the speech processing requirements and tactile channel limitations outlined in Chapter 2, there is sufficient evidence to test the PCA algorithm.

It is agreed by all the previous researchers in the tactile aid field that long-term testing and training is needed to properly evaluate a system. Users should wear the tactile aid as much as possible, if not continuously, to facilitate the learning process. This is particularly important for young children, since hearing children learn to interpret the auditory signals and to produce speech at a young age. Therefore an implementation that will allow field trials and some wearability of the tactile aid, will be best suited to determining the effectiveness of the PC analysis

and, indeed, any speech processing algorithm. Additionally, the implementation should allow modifications to the algorithm, as well as customization for any specific requirements of different users. A compact, real-time system implementation allows the system to go to the user for trials and training, rather than forcing potential users to visit the research laboratory, thereby limiting training time.

The system required to implement the PCA algorithm is shown in Figure 15 at a functional level. There are six operations (functional modules) necessary. First, the speech signal must be converted to an electrical signal and this is done by the microphone module. Next the electrical speech signal must be conditioned for use by the signal processing block. The conversion of the speech signal from the temporal to spectral domain, as well as the principal components analysis, are performed in the Signal Processing block.

Once the principal component values are calculated they must be converted into a form suitable for the output or display device(s). This is done by the Output Interface block, in Figure 15. The signals from the Output Interface block then drive an array of transducers, which may be vibro- or electrotactile types. Finally, a power supply must be provided. This may be a commercial unit operating from an AC outlet, for laboratory testing, but the wearable versions will require battery power and appropriate voltage conversions if needed.

Figure 15 Functional Module Diagram

The following sections will discuss the design, implementation and testing of hardware and software required to provide the function of each block in Figure 15. Section 4.2 provides an overview of the hardware alternatives, the choices made and the actual implementations. Section 4.3 discusses the necessary software and the various design decisions involved. The integration of hardware and software, as well as tactile channel testing, is covered in Section 4.4. Sample system outputs are also reported in Section 4.4.

4.2 Hardware

The first hardware design step was to decide whether to implement the signal processing block of Figure 15 entirely by analog electronics or to employ digital signal processing hardware and software. Although the PCA algorithms could be

implemented using analog electronics only, there are a number of advantages to having programmability in the system.

The implementation requirements of the PCA algorithm are fairly well defined. However, factors such as the actual PC basis vector weights, spectral energy scaling and the number of PCs to calculate may require some defining and experimentation. This would be easier if the PCA calculations were programmable and hence could be altered by a software change. Thus, the spectral analysis could be performed via analog electronics and the PCs calculated digitally.

An all-digital system would allow more extensive variations of the algorithm to be tested. The characteristics of the signals output to the tactile channel could be optimized for each user, as Blamey reported was necessary for some types of tactile displays (Blamey and Clark, 1987). A fully programmable system would also allow entirely different algorithms to be implemented in real-time and tested for tactile aids, as well as cochlear implant systems. In a programmable approach the utility of the hardware is increased and the development, testing and evaluation of speech processing algorithms could be expedited. Therefore, a fully programmable implementation of the signal processing block was chosen.

Previous research has demonstrated the importance of training on the performance of hearing impaired subjects with a tactile aid. Thus, the system should be designed to allow maximum access by potential users. The hardware should be

"stand-alone", compact enough to be portable and ideally should be wearable. Bringing the system to the user will allow training in their environment. A portable device will enable the user to wear the aid for longer periods of time than a system constrained to the laboratory would allow. If a wearable implementation is not achieved in the first iteration of the system, the hardware design should allow this iteration to proceed at a later date.

To provide maximum flexibility in the use of the digital signal processing hardware the analog input/output (I/O) blocks of Figure 15 will be implemented on a "daughter" circuit board that will connect to the programmable digital signal processing (DSP) board. Thus, the application specific I/O will be independent of the signal processing block. (The digital and analog hardware designs and test results are discussed in Sections 4.2.1 and 4.2.2, respectively.)

4.2.1 Digital Hardware

The signal processing requirements of the tactile aid suggest that a microprocessor specifically designed for this task be selected for the computing engine. There are a number of so-called digital signal processing (DSP) microprocessors, with a variety of characteristics, commercially available. The most obvious requirement of the DSP processor is that it have sufficient power to implement the PCA algorithm in real-time. The portability requirement of the tactile aid necessitates the use of a CMOS technology part due to the low

power consumption of integrated circuits manufactured in this technology. In addition, the processor should be I/O independent and expandable in terms of memory addressing so that the signal processing electronics may be used for other algorithms and/or applications. The last selection criterion is vendor specific and involves availability, technical support and debugging tools for the DSP microprocessor.

The field of DSP integrated circuits (ICs) has undergone tremendous growth recently and as a result there are a number of fixed and floating-point choices available. These include the TMS320 series from Texas Instruments, Motorola's 56000/1, the Analog Devices ADSP2100, National Semiconductor LM32900, NEC Electronics uPD77230, Philips PCB5010, AT&T WE DSP16, Thomson TS68930 and Matsushita MN1900. Reviews of these parts have been reported (Marrin, 1986; Electronic Design, 1986; IEEE ASSP, 1986), although the choice must be made taking into account the particular requirements of the tactile aid project.

All of the listed devices are CMOS, but the floating-point devices (uPD77230 and TMS320c30) power consumptions are too large ($> 1.5W$) for portable use and thus were not considered further. In terms of the performance requirements, all but the Thomson and Matsushita offerings have the necessary speed and processing power. This leaves the TMS320c25, 56001, ADSP 2100, National LM32900 and AT&T parts. Of these, the ADSP 2100 and LM32900 both require external data memory, where as the other DSP chips have onboard data

memory. This on-chip memory is large enough for the PCA algorithms and therefore system power consumption and size may be reduced over those parts requiring external RAMs. With the field narrowed to three choices, the final decision was made based on vendor support, part availability, upgrade path and design tools for both hardware and software debugging. The TMS320c25, from Texas Instruments was clearly superior in these respects; at least at the time of the project initiation. Texas Instruments has developed both a software development tool, which was purchased by Carleton University and a full in-circuit emulator, which was made available by the TI Regional Technology Centre in Ottawa. Thus all the necessary development tools were accessible at the start of the project. In addition, the TMS320c25 has been used successfully in other speech processing projects at Carleton University (Morris and Barszczewski, 1989; Morris, 1987).

Microprocessors designed for signal processing applications contain many specialized features. Typically, the instruction set is reduced as compared to standard microprocessors, thus freeing up some chip area for such arithmetic enhancements as a single-cycle multiplier, 32 bit accumulator and registers to store intermediate results. The TMS320c25 includes all of these, as shown in Figure 16. Note also that the TMS chip has a modified "Harvard" architecture; that is, the program and data memories are separate. This

Figure 16 TMS320c25 Architecture (TMS320c25 User's Guide, 1988)

allows data fetches and program instruction execution to occur simultaneously. In addition, there is a large on-chip data memory (544 words) and 4k words of program ROM. Although this program memory must be programmed in volume by Texas Instruments, an EPROM version has been announced (Texas Instruments, 1989) which will allow the user to perform the

programing. In this configuration, as will be seen later, no external memories would be required to implement the PCA algorithms.

The arithmetic unit (ALU), shown in Figure 16, can execute a 16 x 16 bit multiply and a 32 bit accumulation in a single instruction cycle (100 ns). There are also three shift registers to scale data and results as necessary. The instruction set, although not as extensive as that of a standard microprocessor, still provides much flexibility for signal processing algorithms. The instruction format allows operations such as scaling (shifting), counter increments, data moves and pointer adjustments to be done in parallel with other operations. The TMS320c25 also has a versatile I/O setup, with 16 parallel ports and one serial port. Thus the TMS320c25 contains all of the necessary elements to perform efficient signal processing. Additional information may be found in the Texas Instruments Second Generation User's Guide (1988).

With the microprocessor selected the appropriate memories must be specified. The amount of program and data memory storage space is dependant on the algorithms to be implemented. An estimate of the required program memory space may be obtained, given the desired sampling rate of the system, the instruction cycle time and the requirement that the programs execute in real-time. The maximum number of instructions processed between samples, assuming all single-cycle operations and lowest sample frequency of 8 kHz, is:

Sample rate / Instruction cycle
= 125 μ s / 100 ns
= 1250 instructions.

Most TMS320c25 instructions require a single 16-bit word for storage, but some require an extra word for immediate data or address storage. Thus in the worst case, 2500 program memory storage locations will be required. A 4k word program memory should be sufficient for real-time applications, including provision for constant storage and some additional memory if required. This program memory must be external to the TMS320c25, since the on-chip PROM is programmable only by the factory. Note however, that the EPROM version of the TMS320c25, with 4K x 16 program store, (Texas Instruments, 1989) will greatly simplify and miniaturize the hardware for the tactile aid in future iterations of the design.

The data memory requirements are dependant on the algorithms being implemented. In the case of the PCA algorithm, the implementation may be designed to use only the 544 words of on-chip data RAM. To provide some flexibility in the signal processing block, particularly for other algorithms, external data memory is also implemented. The choice of data memory size is dictated largely by the availability of RAMs fast enough to be accessed in a single cycle by the TMS320c25. Most of the fast RAM memory ICs are available in the 4k x 4 bit configuration. Thus, four devices are required to provide 4k words of data storage. Although this would be sufficient in most applications, there is also an 8k x 8 bit RAM available from Performance Semiconductor

Inc., which would allow twice the storage capacity in half the number of parts, with lower power consumption than required by the four 4k x 4 ICs. As illustrated in Figure 17, the signal processing circuit includes the 8k x 8 static

Figure 17 Signal Processing Hardware Blocks

RAMs, the program memories which are 2k x 8 EPROMs from Cypress Semiconductor Inc., address decoding and generation of READY signal for the TMS320c25.

The address decode and READY signal generation circuitry must meet the addressing requirements of the TMS320c25 in terms of timing and memory map organization, as

well as maintaining minimum power consumption and printed circuit board (PCB) size. The TMS320c25 can access 64k words of program and 64k words of data memory. The program memory must be mapped starting at location 0000, since the TMS320c25 interrupt vectors are located here. The external data memory must not overlap the on-chip data memory addresses. The possibility of future memory expansion must also be accounted for when designing the memory maps and decode circuitry.

Figure 18 shows the memory map used in the DSP circuit. Notice that there are memory images of both the program and data memories, but the upper 32k words of both spaces may be used in future memory expansions. The memory images are a

Figure 18 Program and Data Memory Maps

result of partial address decoding which allows the decode circuit to be implemented in only two logic ICs, as illustrated in Figure 19.

Figure 19 Address Decoding Circuit

Due to the fast operating speed of the TMS320c25 microprocessor (100 ns), there are critical timing considerations in the design of memory interface circuitry. (Memory interface timing details may be found in Appendix IIA.) In order to take advantage of the full processing power of the TMS320c25 the program and data memory hardware is designed to be accessed without wait-states. This is

accomplished by the use of fast memory ICs combined with efficient memory address decoding.

As shown in the timing diagrams in Appendix IIA, a READY signal must be provided to the microprocessor for control of memory access cycles. Figure 20 shows the circuit that generates the necessary signal. (Again, the timing details are provided in Appendix IIA.) Note that in Figure 20

Figure 20 READY Signal Generation Circuit

no additional ICs are required because the unused gates of the address decode chips are employed. Thus the two ICs provide both address decoding and READY signal generation. The jumpers in the circuit of Figure 20 allow slower expansion memory and/or I/O devices to be interfaced to the microprocessor, as may be required in other applications.

The signal processing circuitry also requires connectors for the application specific I/O daughter board

and for a memory expansion board. Appendix IIB contains the full digital signal processing circuit with the connectors and the 40 MHz oscillator system clock. The I/O connectors provide all of the signals for the complete I/O space of the TMS320c25 (16 parallel ports, 1 serial port and control signals), thereby allowing maximum flexibility in the input/output circuitry.

The hardware necessary to implement the digital signal processing block of Figure 15 is constructed on a single printed circuit board (PCB). The layout of the PCB was done using the PCB Turbo CAD tool, running on an Intel 80386-based personal computer. For economy of manufacture a two sided circuit board was used, although this complicated the layout of a compact system. (The actual layout may be found in Appendix IIC. Note the density of the tracks required to make the PCB as small as possible. A large prototyping area is provided for experimentation and may be cut off if not required.)

The dimensions of the PCB without the experimentation area are 13.5 x 10 cm. Although this is a prototype layout and includes features that may not be required in a final version, it compares favorably to commercially available speech processors for the deaf. The most popular cochlear implant processor for example, is 13 x 7.5 cm in size (Nucleus Inc., 1987) and employs custom integrated circuits as well as having undergone numerous design iterations.

Appendix IID contains the test results and timing waveforms acquired from the circuit boards. The testing was done using the XDS in-circuit emulator from Texas Instruments. This system allowed full speed hardware testing to be performed under the control of a host personal computer. The programs written to test memory interfacing and accesses are also listed in Appendix IID.

Implementation of the signal processing hardware was achieved on a compact printed circuit board using the TMS320c25 microprocessor, running at a cycle time of 100 ns. The system contains 8k words of data memory and 4k words of program EPROM, expandable to 40k and 36k, respectively. The programmable signal processing hardware allows complex speech processing algorithms to be implemented in real-time. In addition to the principal components analysis algorithm implemented in this thesis, the computationally intensive cepstral analysis formant extraction algorithm (Oppenheim and Schaffer, 1975) has been implemented in real-time on the hardware for use in a cochlear implant system (Gedeon, 1990). Recalling the system functional diagram of Figure 15, input and output interfacing hardware must be designed for the specific application. In this case it is for a tactile speech communication aid.

4.2.2 Analog I/O Hardware

As shown in Figure 15, the signal processing block requires speech signal conditioning as well as an interface

to the tactile transducers. The analog speech signals must be converted to digital form for signal processing by the TMS320c25 and the digital PC outputs from the TMS320c25 must be converted back into analog form to drive the transducers.

At the input to the system, a microphone and amplifier are needed. A popular microphone in tactile aids and cochlear implant speech processors is the Sony ECM-144. This is an efficient, compact lapel microphone that provides flat response over the frequency range of speech. The output signal from the microphone is small (~ 25 mV) and requires an amplifier to boost the signal amplitude. A low-noise JFET operational amplifier (op-amp) is used to match the microphone impedance to the following electronics and to provide gain to the speech signal.

Prior to the analog-to-digital (A/D) conversion, the speech signal must be bandlimited to $1/2$ the sampling frequency ($8 \text{ kHz}/2 = 4 \text{ kHz}$), in order to prevent aliasing of out of band signals. Anti-alias filtering is done by a built-in 8th-order switched capacitor (SC) filter on the A/D converter. A continuous-time filter preceding the SC filter is also included in the A/D device to prevent aliasing in the switched capacitor filter itself.

After gain and bandlimiting stages the speech signal is digitized. The A/D converter selection is determined by three requirements; conversion speed, resolution, and interface characteristics. The sampling frequency chosen for the tactile aid is 8 kHz but the ability to experiment with

higher sampling frequencies up to 16 or 20 kHz may be desired. Thus, the A/D converter must be capable of converting the analog signal at rates up to 20 kHz. In terms of the required resolution, it has been found that 11 or more bits are required in the digital words used to represent the full dynamic range of speech (Witten, 1982). This may seem excessive when one considers that the dynamic range of speech is typically 26-40 dB (Schafer and Rabiner, 1975; Hamsher, 1967) and the approximation that 1 bit represents 6 dB of SNR (Schafer and Rabiner, 1975; Gold and Rader, 1969).

Differences in the average volume over speaker and environments, as well as signal-to-noise requirements in the digital processing operations necessitates the extra digital word length (Schafer and Rabiner, 1975; Witten, 1982). It is also desirable that the A/D converter can be interfaced to the microprocessor with a minimum of additional "glue" logic, in order to keep the circuit size small, reduce power consumption and save I/O time.

The TLC3204x A/D converter, from Texas Instruments, meets all of these requirements and also has a digital-to-analog (D/A) converter built in. Important features of the TLC3204x, also known as the Analog Interface Circuit (AIC), include: CMOS technology fabrication for low power consumption, 14 bit resolution, 19.2 kHz maximum sampling frequency, A/D and D/A on a single IC and a fast serial port interface to the TMS320c25. The TLC3204x also has a programmable gain stage at the input and an auxiliary input

which may be useful in dual microphone noise cancelation schemes in the future. In addition, software selectable antialias and reconstruction filters are provided. A recent version, the TLC32044, has an input antialias filter designed specifically for voice-band applications. The input filter is an 8th-order bandpass with a lower corner frequency of 100 Hz. More information is provided in the TLC320 series data sheets (1989; 1987). As depicted in Figure 21, the interface to the microprocessor does not require any additional logic. Figure 21 shows the portion of the I/O daughter board containing the TLC32044 and TMS320c25 interface.

The many hardware features of the TLC3204x series of analog interface ICs come at the expense of complex software required to initialize the converters and communicate to the microprocessor. The listing of the test program used to evaluate the laboratory TLC32044 setup may be found in Appendix IIE. Note that approximately 200 lines of code are required for initialization and communication, but once set up, only very simple interrupt handling is required.

Automatic gain control (AGC) and noise reduction are not implemented in the present design. Although these two processing steps may be required in a real-world setting, for laboratory and field testing, the speech amplitudes and external noise can be controlled so these steps are not necessary. Noise reduction, in particular, is necessary for most speech processing algorithms associated with tactile aids used in uncontrolled environments.

Figure 21 TLC3204x Interface

The effect of noise on the PCA algorithm is not well documented, but there is limited evidence to suggest that the principal components of speech are fairly robust in the presence of noise (Rajasekaran, and Doddington, 1985; Zahorian, 1988). Also, a noise reduction IC, suited to hearing aid applications, is available commercially (General Instruments, 1988) and may be added if necessary.

At the output end of the system functional diagram in Figure 15 the tactile transducers and interface circuitry are required. As outlined in Chapter 2 there are a number of transducer types and output array arrangements possible. Vibrotactile stimulation was selected over electrotactile stimulation due to the problems of working with direct electrical stimulation. The model V1220 vibrators selected are commercially available from Audiological Engineering Inc. and are used in the Tactaid series of tactile aids, as well as in the latest version of the Queen's University tactile vocoder (Frost, 1988).

The V1220 vibrator is compact (26 x 18 x 10 mm) and lightweight (7 grams). Vibration is produced by a coil of wire on a metal spool that moves vertically between metallic end plates. The motion is coupled to a thermoplastic case which in turn imparts the vibration to the skin surface. The vibrators are designed with a resonant frequency of 250 Hz $\pm 10\%$. Power consumption for these devices is approximately 90 mW for an output amplitude of 40 dB above the perceptual threshold.

The frequency and amplitude response for six individual transducers were tested using the arrangement in Figure 22. The vibrator was strapped on the forearm to provide the correct mechanical coupling. A thin piece of Kynar piezoelectric film placed between the transducer and the arm, sensed the vibration signal. The output signals from the

Kynar, for a swept sinusoidal input are shown in Appendix III for all transducers.

Figure 22 Transducer Evaluation Apparatus

Figure 23 depicts the response of a typical vibrator. Note the sharp resonant frequency peak. The average (across the 6 vibrators tested) frequency of peak vibration amplitude was 258.8 Hz, $\pm 6.2\%$ and the variation of peak amplitude between transducers was less than 1.5 % of the peak value. Thus, all of the vibrators had similar characteristics.

It was determined that the static force applied to the vibrator affected the vibration amplitude, both as measured via the Kynar transducer and as perceived by the subject. Therefore a mass that provided a comfortable force was used to load all transducers equally during testing.

Figure 23 Transducer Frequency Response

Having selected and evaluated the vibrotactile transducers, the vibrator interface to the signal processing block was designed. A 250 Hz, square waveform was used for the transducer carrier signal. A square wave carrier was chosen due to the ease of generation via software. The 250 Hz

carrier frequency was chosen for optimum tactile sensitivity. Depending on how the positive and negative values of the principal components are mapped to the transducers, 4 - 7 vibrators will be required. Thus the interface circuit must be able to multiplex the single D/A output to the appropriate transducer in the array. The current requirements of the vibrators must also be matched by the interface amplifiers.

Figure 24 depicts the required interface hardware. Only one transducer driver is shown in Figure 24, for clarity. (The complete schematic may be found in Appendix IIF.) As shown, the analog signal from the D/A chip passes to the DG201 analog switch IC (two ICs are used for 8 transducers). A data word, output by the TMS320c25 via a parallel port, is decoded by the 74HCT534 and controls the DG201 switches. In addition, the external flag signal, XF, from the microprocessor, can enable or disable all vibrators.

A low-power CMOS opamp, combined with a Class B transistor output driver (Sedra and Smith, 1982) provides the current required by the transducer. A "hold" function is implemented with C1 in order to maintain the vibrator signal level until the next output sample from the D/A arrives. The analog electronics were combined on a daughter circuit board that connects to the DSP mother board. The analog circuit board is 9 x 6 cm in size.

energy storage capacity of 600 mAh, but a nominal voltage of 2.9 V (MoliEnergy, 1989).

The lithium rechargeable cells have a number of advantages over more common Ni-Cad (Nickel-Cadmium) types. Specifically the lithium batteries offer higher energy capacity for a given size, much better charge retention, capability to provide state-of-charge information, higher nominal voltage (hence fewer cells in series are required) and more stable voltage at the end of the discharge cycle. The power consumption of the wearable version and typical battery life are discussed in Section 4.4 System Integration.

4.3 System Software

The software development for the TMS320c25 was done using the Software Development System (SWDS) simulator, assembler/linker and analog interface board (AIB), all from Texas Instruments. The SWDS is a IBM-PC card containing a TMS320c25 and RAM, coupled with user interface software running on the host computer. The comprehensive software debugging features of the SWDS system, when combined with the A/D and D/A of the analog interface board, provided an efficient environment for real-time software development. Detailed information on these design tools may be found in the SWDS User's Guide (1988) and Analog Interface Board User's Guide (1987).

Figure 25 illustrates the functional block diagram of the system software. The software must handle I/O tasks and

signal processing tasks in order to calculate the speech principal spectral components. The I/O task software is essentially the same as the code used in the testing of the I/O hardware, discussed previously. The signal processing modules necessary to implement the principal components analysis include spectral analysis, energy detection and nonlinear amplitude scaling and finally the principal components calculations. These three functional modules are discussed separately.

Figure 25 Software Functional Block Diagram

4.3.1 Spectral Analysis

As outlined in Chapter 3, the PCA algorithm is applied to the incoming, instantaneous speech spectrum. Thus, the

first step in calculating the PCs involves separating the complex speech signal into frequency bands. The characteristics of the frequency bands, centre frequency, bandwidth and out-of-band attenuation, are chosen to approximate the spectral analysis performed by the human auditory system (Searle et al., 1979).

The speech spectrum may be calculated digitally by a number of means including the Fourier transform, (usually via the efficient fast Fourier transform- FFT), linear predictive coding (LPC) or by a bank of bandpass filters. The first two methods, the FFT and LPC, typically operate on a frame of digitized speech of 10 - 32 ms in length, whereas the filter bank processes the speech on a sample by sample basis. Due to the large computational requirements of the FFT and LPC methods, the speech frame length is usually close to 30 ms (at 8 kHz sample rate), in order for the algorithms to be performed in real-time.

Although some speech sounds remain stable over this time, Searle et al. (1979), have found that high frequency components of stop consonant sounds may change up to 40 dB in less than 3 ms. Thus the FFT and LPC schemes that process a window of speech data tend to smear these rapid temporal changes in the speech spectrum. The importance of temporal resolution in the high frequency range of speech has been demonstrated in the stop consonant recognition experiments of Searle et al. (1979), the auditory nerve response experiments of Kiang and Moxon (1974) and the speech perception

experiments of Blumstein and her colleagues (Behrens and Blumstein, 1988; Stevens and Blumstein, 1981; Blumstein and Stevens, 1979).

As discussed in Section 3.3.1, the filter bank method of spectral analysis may be configured to provide rapid temporal response in the high frequency bands by employing filter bandwidths that are a proportion of the filter centre frequency. In the case of the human auditory system, this proportion is 1/3 of an octave, for bands above 400 Hz (Kiang and Moxon, 1974). Thus bandpass filters provide spectral analysis that best approximates that of the auditory system.

There are two types of digital filters distinguished by the arrangement of feedback. The Infinite Impulse Response (IIR) type has feedback of both past outputs and past inputs, whereas the Finite Impulse Response (FIR) type has feedback of only previous inputs. By definition, the impulse response duration of an FIR system is finite, thus the output may be expressed as a finite convolution sum:

$$y(n) = \sum_{k=0}^M h(k) * x(n-k) \quad (4)$$

where $y(n)$:= output sample at time n
 $h(k)$:= response of the system to unit sample k
 $x(n-k)$:= is the input sample at time $n-k$.

Equation (4) may also be thought of as a weighted moving average of the input samples. Two important characteristics of the FIR digital filters are the inherent stability and the ability to achieve linear phase response. Since FIR filters

do not employ output feedback they cannot be unstable. This is an important factor when implementing filters using fixed-point arithmetic because the limited precision of the fixed-point implementation can result in unstable filters.

IIR filters, as the name implies, can have an infinite impulse response, since previous output values are fed back into the system. Equation (5), is the difference equation for an IIR system.

$$y(n) = - \sum_{k=1}^N a(k) * y(n-k) + \sum_{k=0}^M b(k) * x(n-k) \quad (5)$$

Note that in equation (5), the weighted sum of previous outputs $y(n-k)$, causes the impulse response to theoretically continue forever. This recursive portion may produce instability in the filter output $y(n)$, depending on the values of the weighting function $a(k)$. Instability may be a problem particularly when implementing IIR filters using fixed-point arithmetic. In addition, IIR filters have nonlinear phase response which may be detrimental in some applications.

The advantage of IIR filters over FIR types is that IIR digital filters of a given order (complexity) are directly equivalent in amplitude response to an analog counterpart of the same order (Rabiner and Gold, 1979). For example, a 6th-order Butterworth analog filter will have similar magnitude characteristics as an IIR 6th-order, Butterworth filter. FIR filters require much higher filter order to achieve

equivalent magnitude response, particularly for stopband attenuation. Figure 26 shows two bandpass filters, one a 6th-order IIR design and the other a FIR filter with an impulse response length of 75 samples (75th-order).

Figure 26 FIR vs IIR Magnitude Response

Despite the difficulties involved with IIR digital filter implementation, IIR filters were selected for the present system because they allow more rapid attenuation in the stopbands than would be possible with a real-time FIR filter bank. It should be noted that the TMS320c25 instruction set provides an efficient means of implementing FIR filters via the RPTK repeat instruction, coupled with the

MACD, multiply / accumulate / data move instruction.

Essentially, FIR filters may be executed at a rate of 100 ns per tap, therefore the 75 tap FIR filter of Figure 26 will take 7.5 μ s to execute, whereas the 6th-order IIR filter will take approximately 5 μ s (as will be shown later). Considering that the sample interval is only 125 μ s, the time difference per filter becomes important when it is multiplied by the number of filters in the bank. A FIR filter bank with 12 bandpass filters would require 90 μ s (72% of real-time) to process, while the IIR version would take only 60 μ s (48% of real-time). Moreover, the linear phase characteristics of the FIR filters are not required in the present application since only the frequency band energies are required for the PCA algorithm.

With the filter type (IIR) selected, the filter bank may be designed. As discussed in Chapter 3, filters based on 6th-order Butterworth designs, with 1/3 octave bandwidths, are similar to human auditory tuning curves and are well suited to PCA of speech (Searle, 1979). Table 4 contains the filter centre frequencies and bandwidths for the tactile aid filter bank. The frequency parameters in Table 4 correspond to the ANSI S1.11-1966 Class III filter specification for 1/3 octave band filter banks. Most commercial filter banks used by research groups in PCA experiments, are designed to meet the ANSI standard; hence the spectral analysis performed for the tactile aid will be similar to past research (Brooks and Frost, 1983; Bustamante, 1986; Klein, Plomp and Pols, 1970,

Searle, 1985). Note that the filter bandwidths below 500 Hz are 2/3 octave wide as per the work of Searle et al. (1979) outlined in Chapter 3.

Filter Number	Centre Frequency (Hz)	Bandwidth Response (Hz)	Impulse Response (ms)
14	3250	732	1.2
13	2500	582	1.8
12	2000	462	2.0
11	1600	367	2.5
10	1250	291	2.6
9	1000	231	2.9
8	800	184	3.7
7	630	143	4.6
6	500	116	6.0
5	400	92	6.5
4-3	282.5	131	7.0
2-1	180	83	7.5

Table 5 Filter Bank Centre Frequencies and Bandwidths

The individual filters were designed using the Digital Filter Design Program (DFDP) from Atlanta Signal Processors Inc. (1986) combined with a Fortran program (Parks and Burrus, 1987) that was modified slightly for use in the present application (see Appendix IV-A). The Fortran program was required because LFDP uses frequency and attenuation values as input to the design process and does not allow the filter order to be specified. Thus, the Fortran program was used for initial floating-point filter designs and the filter frequency response curves provided the necessary input

parameters to DFDP in order to produce filters of the desired length. DFDP was also used to simulate the effects of filter coefficient quantization in 16 bit fixed-point simulations and to generate plots of filter prototype performance.

Figure 27 shows a DFDP plot of the filter magnitude vs. frequency for the bandpass filter with a centre frequency of 1 kHz from the filter bank. Also shown in Figure 27, is the ANSI S1.11-1966 standard Class III 1/3 octave filter

Figure 27 DFDP CH9 Magnitude Plot and ANSI 1966 S1.11 Specification

requirements. Although exact compliance with all aspects of the ANSI standard would require considerable effort in fine-tuning, the filters are designed to be within 0.5 dB of the ANSI minimum attenuation characteristics for all frequencies.

Given the transfer function of a digital filter, there are a very large number of difference equations that may be formed. For example, the IIR filter difference equation (5) above, is a direct implementation of the filter transfer function:

$$H(z)=B(z)/A(z)= \frac{b_0 + b_1 z^{-1} + b_2 z^{-2} + \dots + b_M z^{-M}}{1 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_N z^{-N}} \quad (6)$$

The digital filter difference equation may be represented graphically as a signal flow diagram. The graphical representation can a useful aid in analyzing the characteristics of various difference equations. The direct form IIR signal flow diagram, corresponding to equation (5), is shown in Figure 28, for a second-order filter.

Depending on the IIR filter implementation structure (difference equation form) used, fixed-point implementation problems may occur. These implementation problems include:

- i) filter coefficient errors due to quantization,
- ii) quantization noise and overflow from arithmetic operations,
- iii) low amplitude oscillations (small-scale limit cycles) caused by quantization in the feedback loop and
- iv) large-scale limit cycle oscillations caused by overflow in the feedback loop.

Figure 28 Direct Form IIR Signal Flow Diagram

The first two problems change the filter frequency response while the oscillation problems produce undesirable output variations. In particular, the limit cycle oscillations have amplitude variations encompassing the full dynamic range of the filter and, therefore, cannot be tolerated.

The fixed-point problems are difficult to analyze for IIR filters due to the number of feedback loops. It has been found, however, that various implementations of the same filter transfer function have different sensitivities to the fixed-point problems. For example, the coefficient sensitivity problem and its relation to the filter structure has been examined by Erickson and Michel (1985) and Kaiser (1965) and reported in Parks and Burrus (1987). The IIR filter transfer function $H(z)$ in equation (6) may be expressed in terms of positive powers of z as:

$$H(z) = \frac{z^{N-M} (b_0 z^M + b_1 z^{M-1} + \dots + b_M)}{z^N + a_1 z^{N-1} + a_2 z^{N-2} + \dots + a_N} \quad (7)$$

$$H(z) = z^{N-M} \frac{B(z)}{A(z)} \quad (8)$$

To examine the effects of quantizing the coefficients, the variation in location of the filter poles in (7), as a function of changes in the coefficients a_k is investigated. As per Parks and Burrus (1987), the denominator polynomial in (8), may be expressed as a Taylor series expansion in terms of z and a_k :

$$A(z_m + \Delta z_m, a_k + \Delta a_k) = a(z_m, a_k) + \Delta a_k \frac{\partial A(z)}{\partial a_k} + \Delta z_m \frac{\partial A(z)}{\partial z_m} + \dots$$

After some manipulations, the change in pole location, Δz_m , may be expressed in terms of the change in filter coefficients, Δa_k , as in equation (9):

$$\Delta z_m = \Delta a_k \left| \frac{z_m^{N-k}}{\prod_{j=1; j \neq m} (z_m - z_j)} \right| \quad (9)$$

From equation (9) it can be seen that the coefficient sensitivity increases when the poles are close together, since $z_m - z_j$ will be small. Also the filter will be most sensitive to the last coefficient a_N because $N-k$ will be zero. Moreover, the sharper the attenuation of the filter the closer the poles will be to the unit circle, (where $|z|=1$) and the more sensitive the pole location will be to coefficient variation, since the numerator of (9) will be large.

Filters with sharp transition regions are particularly prone to coefficient sensitivity problems because the filter order, N , is high which makes it difficult to separate the poles and the poles tend to be near the unit circle. The solution to this problem is the use of multiple lower order filter sections to realize the higher order filter function. This allows the poles in the each low order section to be separated, thereby reducing the pole location sensitivities to coefficient quantization. Typically second-order blocks are used to implement higher order IIR filters.

There are many possible arrangements of the second-order blocks, as well as different forms for the blocks themselves. In terms of the overall filter structure, there are two common arrangements for the second-order blocks, as shown in Figures 29 and 30. Figure 29 shows the cascade structure, with the low order blocks placed in series. As depicted in Figure 30, the second structure uses a parallel arrangement of second-order blocks. The filter transfer

function forms for the cascade and parallel second-order implementations are listed below, as equations (10) and (11), respectively.

Figure 29 Cascade Structure

The corresponding cascade transfer function is:

$$H(z) = K * \prod_{k=1}^3 \frac{1 + b_{1k}z^{-1} + b_{2k}z^{-2}}{1 - a_{1k}z^{-1} - a_{2k}z^{-2}} \quad (10)$$

Figure 30 Parallel Structure

The parallel structure transfer function can be found via partial fraction expansion (Oppenheim and Schafer, 1975) and is:

$$H(z) = \sum_{k=1}^3 H_k(z) + P_k ;$$

(11)

$$\text{where: } H_k(z) = \frac{b_{0k} + b_{1k}z^{-1}}{1 - a_{1k}z^{-1} - a_{2k}z^{-2}}$$

In the cascade scheme, many different structures may be designed, simply by varying the order of the blocks and the pole-zero pairings of the overall transfer function. It is also possible to use different implementation structures for each second-order section. Thus, sections with special requirements, such as poles near the unit circle, may be implemented in the structure with the most desirable characteristics. Another important advantage of the cascade implementation is that zeros, located on the unit circle ($z=1$) are easily implemented (Jackson, 1970; Parks and Burrus, 1987). In the numerator of equation (10), if a single second-order section has a unit circle zero, then the overall filter will also have a zero at $z=1$.

In the parallel structure, the unit circle zeros are more difficult to implement, since the zeros depend on

cancellation of terms in equation (11). Moreover, the zeros are more sensitive to coefficient quantization in the parallel structure (Parks and Burrus, 1987). The parallel structure however, does have the advantage that the scaling of each second-order block is independent of the others, unlike the cascade structure. Also the choice of cascade order is not required in the parallel structure.

Due to the coefficient sensitivity of the parallel structure, the cascade of second-order blocks is used for all filters in the present application. The pole-pairing and order of the second-order blocks in the cascade was done according to the guidelines proposed by Jackson (1970a; 1970b). To reduce the filter noise due to rounding and scaling as required in fixed-point arithmetic, while also preventing overflow for narrowband signals, Jackson developed the following rules:

- 1) The complex pole pair nearest the unit circle (highest Q factor) along with the zero nearest to it, are implemented in the last second-order block.
- 2) The other poles and zeros are paired similarly, thus resulting in a set of second-order blocks with increasing Q .
- 3) The gain of the filter is divided among the blocks so that the maximum gain of any single second-order block is one.

These rules, although conservative, do provide good filter responses and are applied to all the filters in the present application.

The filter form used for the individual second-order blocks is derived from the direct form, shown in Figure 28,

as follows. Note that the direct form, illustrated in Figure 28, requires four delay storage registers per second-order section. These may be reduced to two delay registers as shown in Figure 31. The same filter transfer function may be implemented by both second-order blocks, but the difference equations are dissimilar.

The difference equation corresponding to the filter in Figure 31 is:

$$y(n) = \sum_{k=0}^M b_k w(n-k) ; \quad (12)$$

$$\text{where } w(n) = x(n) - \sum_{k=1}^M a_k w(n-k) .$$

Figure 31 Direct Form II Second-order Filter

Note the difference from equation (5) p.126, for the direct form requiring four delay storage locations. Both direct filter forms require the same number of mathematical operations, but the direct form II requires less memory for delay value storage.

Although the direct form II configuration is common in floating-point filter implementations, Jackson (1970) found that a transposed configuration was less susceptible to the problems of roundoff noise and dynamic range limitations and thus better suited to fixed-point systems. Application of the transposition theorem for linear systems (Oppenheim and Schafer, 1975) to the signal flow graph in Figure 31 leads to the transpose form as used in the bandpass filter bank for the PCA algorithm. Essentially, the theorem states that when all signal paths are reversed the input and output are reversed, all summing junctions are changed to branching nodes and, conversely, all nodes become summing junctions. The resulting signal flow diagram is shown in Figure 32. Mathematically, this process is carried out by determining the transpose of the discrete-time transfer function matrix, hence the name 'transpose form' (Jackson, 1970).

Figure 32 Transpose Form of a Second-Order Filter

Figure 33 shows the TMS320c25 assembly language code of first and last second-order sections used to implement a 6th-order, IIR bandpass filter in a cascade of three second-order sections of the transpose form. The code listed is for the bandpass filter with centre frequency of 1 kHz. Note that the input to the filter is scaled and that the scaling in each second-order section may be different in order to optimize the conversion of the filter coefficients into fixed-point representations. Note, also, that the coefficients are accessed using the direct addressing mode of the TMS320c25. If the coefficient is small enough (less than 13 bits) it is placed directly within the multiply instruction, which saves memory.

The code for the other filters in the system is similar to that in Figure 33. The filter bank program uses an inline format of filter code, to eliminate looping overhead. The

actual frequency response for the 1 kHz bandpass filter is shown in Figure 34, along with the other two filters that make up the octave. This response was measured by digitizing the swept sinusoid test signal from a spectrum analyzer, filtering in real-time using the TMS320c25 and converting back to the analog domain, to input to the spectrum analyzer. For the plots in Figure 34, the 12-bit A/D - D/A of the Analog Interface Board was used, hence the approximate 72 dB noise floor, as indicated on the figure.

```

.....
*   FILTER GENERATED FROM FILE
*   ch9i68.fl
*   .....
*   OUTPUT: VCH9i6
*   .....
*
* SECOND-ORDER FILTER SECTION # 01
*
  LAC   INPUT,13      *GET & SCALE INPUT
  SACH  FILTT         *SAVE SCALED INPUT
  LT    FILTT         *GET SCALED INPUT
  MPY   B910          *P = B0 * INPUT
  ZALH  Z911          *AC = Z-1
  APAC                      *AC = Z-1 + (B0 * INPUT)
  SACH  VCH9i6        *Save in OUTPUT
  LT    VCH9i6        *Get OUTPUT
  ZALH  Z912          *AC = Z-2
  MPY   A911          *P = A1 * OUTPUT
  APAC                      *AC = (B1 * INPUT) + (A1 * OUTPUT)
  MPYA  A912          *P = A2 * OUTPUT
                      *Save in Z-1
  SACH  Z911          *AC = A2 * OUTPUT
  LTP   FILTT         *P = B2 * INPUT
  MPY   B912          *AC = (B2 * INPUT) + (A2 * OUTPUT)
  APAC                      *Save in Z-2
  SACH  Z912
*
* SECOND-ORDER FILTER SECTION # 02
*
* SECOND-ORDER FILTER SECTION # 03
*
  LAC   VCH9i6,14     *GET & SCALE INPUT
  SACH  FILTT         *SAVE SCALED INPUT
  LT    FILTT         *GET SCALED INPUT
  MPY   B930          *P = B0 * INPUT
  ZALH  Z931          *AC = Z-1
  APAC                      *AC = Z-1 + (B0 * INPUT)
  SACH  VCH9i6        *Save in OUTPUT
  LT    VCH9i6        *Get OUTPUT
  ZALH  Z932          *AC = Z-2
  MPY   A931          *P = A1 * OUTPUT
  APAC                      *AC = (B1 * INPUT) + (A1 * OUTPUT)
  MPYA  A932          *P = A2 * OUTPUT
                      *Save in Z-1
  SACH  Z931          *AC = A2 * OUTPUT
  LTP   FILTT         *P = B2 * INPUT
  MPY   B932          *AC = (B2 * INPUT) + (A2 * OUTPUT)
  APAC                      *Save in Z-2
  SACH  Z932
*

```

Figure 33 TMS320c25 Program for an IIR Bandpass Filter

Figure 34 1 kHz Octave Frequency Response

Plots of the other filters in the bank may be found in Appendix IV-B. Note that all filters are within 0.5 dB of the ANSI 1966 specification for 1/3 octave filters and the lower two channels, with 2/3 octave bandwidth were increased to 8th-order to better comply with the ANSI magnitude specification. The variation in passband response between filters in the bank is less than 0.2 dB. The in-band ripple is less than 0.1 dB in all cases. Table 5 contains the program and data memory space requirements and speed of execution characteristics of the filter bank.

Number of filters	12
Execution time per filter Total time	5.1 μ s (6th-order) 64.6 μ s
Memory per filter	Data - 20 words Program - 51 words
Total Memory	Data - 252 words Program - 646 words

Table 6 Filter Bank Characteristics

4.3.2 Spectral Energy Detection and Conditioning

As shown in the software block diagram of Figure 25, the energy level in each bandpass filter must be determined and then nonlinearly scaled. This processing is in keeping with the desire to simulate the human auditory system. Moreover, the nonlinear scaling scheme and energy detection are chosen to emulate the analog systems used for principal components analysis by previous researchers.

In analog electronic hardware schemes, the energy detector usually consists of a fullwave rectifier and lowpass filter to convert the sinusoidal-type outputs of each filter, into a time-varying DC value (Searle, 1985; Brooks, 1984). Figure 35 illustrates analog type of signal processing and the waveforms at each step. Note that depending on the LPF time constant, the end result is approximately the peak value of the filter output signal.

Figure 35 Analog Energy Detector Scheme

Emulating the analog process shown in Figure 35 using digital signal processing methods poses a problem in terms of the required sample rate. For example, the fullwave rectification step is easily done in software via the ABS (absolute value) instruction. However, the resultant signal contains significant energy at twice the frequency of the original. In order to prevent aliasing in the digital lowpass filter that follows, the sample rate must be increased to at least twice the new frequency (4 times the original sampling frequency).

Since the end result of the analog energy detector is approximately equal to the peak value of the original bandpass filter output (see Figure 35), a simple peak detecting scheme is used in the present system. In a frame of filter output values the largest amplitude output sample is retained. At the end of the frame this value is then used in the principal components calculation. The frame lengths are chosen, in combination with the filter rise-times, to give a filter/detector rise-time similar to the human auditory system as per Searle's work (Searle, 1985; Searle et al.,

1979; Searle et al., 1978). Table 4 contains the filter/detector impulse rise-times.

Typical response to a suddenly applied sinusoid, at the centre frequency of the bandpass filter, is shown in Figure 36, for the 1 kHz bandpass filter/detector combination. Appendix IV-C contains the TMS320c25 code that performs the peak-picking function for a single channel. Total execution time of the peak-picking code for all 12 channels is 19.2 μ s (max), with 36 data memory locations and 202 words of program storage required.

As outlined in Chapter 3, Zahorian (Zahorian, 1988; Zahorian and Rothenberg, 1981; Zahorian and Jagharghi, 1985) found that the principal components analysis speech coders performed best when the spectral outputs were scaled logarithmically. Log scaling is also used by Searle (1985) to emulate the psychophysical performance of the human auditory system, in terms of the wide dynamic range of human hearing. Moreover, Brooks (1984) determined that log scaling of the filter energy in a tactile vocoder provided the optimum performance in phoneme and word recognition tasks. Therefore logarithmic scaling of the filter outputs is performed in the present system.

Figure 36 Channel 9 Filter/Detector Response to Sudden Sinusoid Input

Unfortunately implementation of the log function on the TMS320c25 is difficult. A 'log' instruction is not provided in the instruction set, thus an approximation to the logarithm must be implemented with the available instructions. Even this is not straightforward on the TMS320c25, due to the time constraints and fixed-point arithmetic. In a similar project, (Engebretson and O'Connell, 1986) developed a log approximation for the TMS32020 DSP that required 26 μ s to execute. Scaled to the faster processing time of the TMS320c25, this code would require 13 μ s per filter. This is unacceptable since the total time necessary would be $(12 \times 13) = 156 \mu$ s and the sample interval is only 125 μ s. Although it may not be necessary to calculate all filter output

logarithms on a sample by sample basis, the time required per channel would remain prohibitively slow.

A faster method of determining the approximate logarithm of a number is by the use of a piece-wise linear approximation of the log function. In this scheme, a number of linear sections are used to approximate the log function over the range of interest. Calculation of the logarithm of a particular number is done by determining the linear segment corresponding to the number and then solving the equation $y = mx + b$, where x is the input, m and b are stored parameters of the linear segment and y is the log approximation of x .

A Fortran program (LOGAP2.FOR) was written (see Appendix IV-D) to facilitate the selection of the linear segments in the logarithm (base 10) approximation. The LOGAP2 program determines the fewest number of linear segments required, for a given allowable error between the approximation and the actual logarithm function. The program inputs are the starting linear input value, the maximum error and the starting step size to use for the iterations. Additionally, the program interactively allows the user to vary the step size to best suit the linear input magnitude range and to input approximations of the midpoint value on a linear segment, in order to speed the calculations. The program prints the number of linear sections required to meet the specifications, the slope and y-intercepts and the end points for the sections. (The reader is referred to Appendix IV-D for specific details.)

Using LOGAP2, a 12 segment approximation spanning the linear range of 1 to 40000, or 92 dB, with a maximum error of ± 0.5 dB was designed. Figure 37 illustrates the log function and the approximation designed via the LOGAP2 program over the entire input range. The actual log function curve is shifted up +5 dB, in order to allow the detail of the approximation curve to be visible. Figure 38 shows an expanded scale spanning the linear range between 1 and 28. Note that the maximum error can be clearly seen in Figure 38 to be 0.5 dB.

Figure 37 $20(\log(x))$ and 12 Segment Linear Approximation

The slope (m) and intercept (b) floating-point values for the 12 linear segments were converted to fixed-point and scaled as required. As can be seen in Figure 38 the first two linear segments represent the input ranges 1-2 and 3-4. Rather than calculate the linear approximation for the first four data points the actual fixed-point representation of $20 \cdot \log(x)$ was stored for each of the four x values from 1 to 4.

Figure 38 Log Function and Linear Approximation Expanded Scale

A successive approximation search (coded in TMS320c25 assembly language) was used to select the appropriate linear segment when performing a linear to logarithm conversion. Once the correct linear segment is selected the log value is

calculated using the stored slope (m) and intercept (b) values in the formula: $y = mx + b$. Figure 39 plots the TMS320c25 implementation output of the log approximation and the error, both versus the actual floating-point log function. The log approximation code is listed in Appendix IV-D. The worst-case execution time for the log approximation code is 3.2 μ s and the program requires 25 words of data memory and 150 locations of program storage.

Figure 39 TMS320c25 Log Approximation and Error vs the Floating-point Log Function

4.3.3 Principal Components Calculation

The final step in the speech processing chain is the weighted summation of spectral energies to obtain the principal components (refer back to Figure 25). As outlined in Section 3.3.2 of this document, a $1/2$ cosine series weighting function approximation to the principal component weights was used. The floating-point values for the weights were converted to fixed-point representation and scaled to prevent overflow in the summation of bandpass filter energies.

4.3.4 Software Module Integration and Statistics

After the software modules of Figure 25 were implemented and tested individually, they were integrated into a single program. The combined execution time of all software modules, for all channels, i.e. 12 filters, 12 energy detectors, 12 log calculations and the four principal component summations, yielded an execution time of approximately 130 μ s. Given the sample rate of 125 μ s, clearly execution of all software tasks, during every sample period, is not possible. As Searle's work (Searle et al., 1979; Searle, 1985), has demonstrated, the high frequency filter outputs should be sampled at the rate of 1 or 2 ms. For a 1 ms update rate, the calculation of the 12 log values and four PC summations may be distributed over the $(1\text{ms} / 125\mu\text{s}) = 8$ sample periods available. Figure 40 illustrates the division of software tasks over the 1 ms time period.

Note that all filters are executed on every sample but only two log calculations are done per 125 μ s period. Thus the total program execution time per sample is within the allowed limit of 125 μ s.

Note: All filters and peak-pickers are executed on every sample

Figure 40 Software Task Distribution

The performance statistics for the entire principal components program are summarized in Table 6. The execution time and memory requirements for each section of the program are listed, along with the total values. The longest program execution time (BPFs+Peak Pick+PC Sum) occupies 78% of the real-time period available.

Program Section	Memory Space (words)		Execution Time (μ s)
	Program	Data	
BPFs	646	252	64.6
Peak Pick	202	36	19.2 (max)
Logarithm	150	37	3.2 (max)
PC Sum	136	33	14.1 (max)

Table 7 Program Statistics

4.4 System Integration and Testing

Once the principal components spectral analysis software was developed the software and hardware blocks were integrated and tested. The tactile vibrator amplitude levels were scaled according to the informal perceptual tests as discussed in Section 4.4.1. The generation of the tactile stimulus carrier signal and integration of I/O hardware and software are outlined in Section 4.4.2. And finally, some sample system responses to speech input are provided in Section 4.4.3.

4.4.1 Tactile Psychophysical Testing

Detailed psychophysical testing of the tactile sense is beyond the scope of this thesis. However, preliminary evaluations of tactile amplitude perception with the V1220 vibrator were performed. The results of these tests are necessary to select the output signal voltage range for the tactile aid. The test apparatus, procedure and results are

discussed in this section and in Appendix V. Chapter 5 includes an outline of future work to be done in the area of tactile psychophysical testing. Blamey et al. (1987), present psychophysical data obtained in a detailed analysis of their electrotactile display and is a good example of the type of evaluations that could be done with a V1220 vibrator display.

The threshold of amplitude perception and maximum comfortable loudness (amplitude) parameters were tested for three subjects. These two parameters were examined because they have an immediate effect on the hardware and software of the tactile aid. Figure 41 illustrates the apparatus used for the experiment. The tests were driven from a Digital Equipment Corp. (DEC) minicomputer, that included A/D and IEEE488 bus peripherals. Software (see App. V), running on the MINC-11, controlled the amplitude of a 250 Hz waveform via the IEEE488 bus connected to a Wavetek 175 programmable waveform generator.

A method of adjustment procedure was used for the experiments. The subject sat in an isolation booth and adjusted a potentiometer to set the vibrator sensation to threshold, or maximum comfortable level, depending on the test. The potentiometer voltage was digitized and used to control the vibrator signal amplitude. The starting amplitude level was randomized for each trial (i.e. it could start above or below threshold, by a random amount). Additional details of the test programs may be found in Appendix V.

The three subjects tested were TJR (the author) a 26 year old male, NTT a 10 year old male and SET a 12 year old female. All tests were performed using the same V1220 vibrator (Audiological Engineering), worn on the upper side of the left wrist, as one would wear a wristwatch. The threshold and maximum amplitude (measured in Vpp of the input signal) results are plotted in Figures 42 and 43, respectively. TJR had ~ 6 hours practice with the vibrators and was considered an experienced subject. The children, NTT and SET were both novice subjects and had ~ 10 minutes practice with the vibrator set-up before their sessions. Subjects had no known hearing deficiencies.

Figure 41 Tactile Amplitude Perception Test Apparatus

Figure 42 Tactile Amplitude Threshold Results

Figure 43 Maximum Comfortable Level Results

Similar performance was observed for the both children. TJR's thresholds were slightly lower than the children, possibly due to his previous experience, both as a subject in general and with the current tests in particular. In more formal evaluations, Bernstein et al. (1986) found that children and adults had similar thresholds independent of the subject's sex.

As shown in the Figure 42, performance on the threshold tests improved for all subjects as they gained experience. TJR performed more sessions than the other subjects and had more experience with the threshold experiment. His threshold changes can be seen to stabilize at an input signal level of approximately 0.04 Vpp, for the last three sessions.

The maximum comfortable loudness values shown in Figure 43 were the average of 10 trials by each subject. Notice that highest maximum comfortable loudness values were reported by the subject with the most exposure to the vibrations. This was also found by Blamey et al. (1987) for electrotactile stimulation. It is expected that the maximum amplitude values would increase somewhat with subject experience but would remain within the vibrator manufacturer's recommended maximum input level of 8 Vpp. Informal tests near the maximum allowed signal level, indicate that extended use at this level would not be comfortable. Additionally, the manufacturer has determined a signal of 4 Vpp corresponds to a perceptual level of 40 dB above threshold; therefore, input signals at

the maximum amplitude are not likely to be required for any significant length of time.

Additional trials were performed by subject TJR, with the location of the vibrator being varied for the threshold tests. The vibrator was worn on the upper, lower and left and right sides of both the left and right wrists. The average threshold values for all positions was 0.0306 ± 0.011 Vpp. The four positions around the wrist could be used in the final tactile aid display, either on a single wrist, or on both, depending on the number of vibrators used.

The subject average threshold value was 0.14 Vpp, while the average maximum comfortable level was found to be 4.95 Vpp. These data were used in the TMS320 program to scale the PC output amplitude ranges. Additionally, the maximum comfortable loudness value indicates that a ± 5 V power supply for the analog circuitry is sufficient.

4.4.2 Hardware / Software Integration

The software development for the PCA speech processor was done on the IBM PC compatible based Software Development System (SWDS), for the TMS320c25, with the Analog Interface Board (AIB) used for I/O. The final stage in the tactile aid implementation involves porting this software to the testbed hardware for use as a self-contained system, as well as generating the tactile vibrator carrier signal.

The carrier signals used to drive the tactile vibrators were digitally generated, 250 Hz square waveforms, with a

sampling frequency of 1 kHz. The software task distribution diagram in Figure 40, shows that each tactile channel gets an output every millisecond (1 kHz frequency). By alternating the polarity of the output signal, every 2 ms, under software control, a 250 Hz signal is generated. The amplitude of the four carrier signals (one per vibrator) is controlled by the amplitude of the calculated principal component value. The maximum and minimum voltage values were set to 8 Vpp and 0.03 Vpp respectively, as determined by the tactile perceptual test results.

Three of the PC outputs (1,2 and 3), may have negative or positive DC amplitude values. Two simple output schemes are proposed for differentiating between the positive and negative values using the amplitude modulated signal driving the vibrators. The first is to simply DC offset the PC sums, so that the most negative value now equates to a value of 0. In this manner, the vibrator in the off state will indicate the presence of a very large negative PC1, PC2 or PC3 output.

The second scheme involves using two vibrators per PC1, 2 or 3 output, with one activated in the case of positive PC output and the second for a negative signal. No additional power is required, since only one of the two vibrators may be on at a time. The disadvantage is that the number of output devices that the user must attend to is larger. This may actually not be a disadvantage since it may prove easier for the user to distinguish distinct patterns with the extra vibrators due to the increased dynamic range available for

the positive and negative PC values and the physical isolation between the positive and negative stimulus locations. Boothroyd's work with tactile pitch displays (Boothroyd, et al., 1988) demonstrated the superiority of encoding information in the tactile spatial dimension via a multichannel display array, over a single channel output device with encoding in the complex stimulus waveform.

Finally, the program was loaded into EPROMs for use in the stand-alone hardware. Because the program memory map of the hardware is not in a continuous block of 4k words, an unconditional branch instruction was included at the end of the first physical memory block (address 2047), so that the program control would jump to the beginning of the second memory block (address 16384).

4.4.3 System Function Verification

The principal component tactile outputs for sample speech inputs are presented in this section to demonstrate the functionality of the hardware and software. (Detailed subjective testing of the tactile aid, as well as algorithm enhancement evaluations are currently underway and are beyond the scope of this thesis.) Chapter 5 outlines some of this work and recommendations for further evaluations of the tactile aid.

The system outputs and speech input plots were taken from the final, real-time tactile aid system. The speech input was spoken by the author and the real-time system

response was captured using a digital storage oscilloscope. Figure 44 shows the PC0 response to the utterance 'adda'. Note how the PC0 signal tracks the amplitude envelope of the speech signal, as outlined in Chapter 3.

Figure 44 PC0 for Utterance 'adda'

The PC1 output signal for the utterance 'sha' is depicted in Figure 45. The plot of PC1, in Figure 45, has not been DC shifted, but was set up for output in the two transducers-per-PC-scheme. Note that for the initial high frequency fricative, the PC is negative and thus the negative vibrator would be active. As the voiced vowel is produced the

major energy content of the speech moves lower in frequency and the PC value becomes positive. In this case the negative vibrator would be inactive and the positive vibrator would become active.

Figure 45 PC1 for the Utterance 'sha'

4.4.4 Final Tactile Aid Considerations

The final tactile aid considerations include the power consumption and the packaging / wearing considerations. Although the total power consumption of the current prototype (~ 0.75W) is likely too high for a practical wearable aid, battery power would be able to provide ~3 hours of use in 'typical' situations. It should be noted that no power

reduction schemes, such as power down during quiet periods, have been employed. This is an area for future investigation. More efficient output devices and the EPROM version of the DSP processor (TMS320E25) will also serve to reduce the power consumption of future prototypes.

The current prototype hardware consists of a 13.5x10 cm printed circuit board (PCB), containing the digital hardware and a 10x6 cm wire-wrapped analog I/O daughter board. This stand-alone, laboratory unit is packaged in a 16x12x8 cm box, with jacks for the +12 and -12 V supplies, input signal (microphone, or recorded source) and vibrator outputs. An input signal gain adjustment is also provided. A PCB version of the I/O board will allow the entire system, including batteries, to be packaged in a wearable container 13.5x10x4 cm in size. The next version prototypes, currently under development, will be ~ 12x6.5x4 cm in size. Further details on this work are presented in Chapter 5.

Chapter 5

5.0 Conclusions

Perhaps the most important aspect in the development of a prosthetic aid for the handicapped, is a thorough understanding of the problem. Therefore, a significant portion of this thesis has been devoted to the problems associated with, and potential aids to, profound deafness. The communication obstacles facing the profoundly deaf individual are formidable. Hearing loss causes a decrease in human information reception and production capabilities. The problems become particularly acute for children with severe hearing loss since the damaged acoustic feedback path inhibits the development of speech production, as well as reading skills.

As one might imagine, there have been many attempts at developing devices to assist the profoundly deaf. Unfortunately the standard amplification type hearing aids are of little use to individuals with profound hearing loss. The most attractive scheme, therefore, has been to bypass the damaged section of the auditory channel by stimulating the auditory nerve directly, or by using the visual or tactile sense. This thesis has described the design of a programmable speech processor suitable for either type of system and the real-time implementation of a tactile "hearing" aid using the principal components analysis of speech spectra algorithm.

A review of past tactile aid work and more general research relating to the effects of profound deafness on speech reception and production provided guidelines for development of an auditory prosthesis. Due to the inter-subject variabilities discovered, particularly in tactile perception, the need for a programmable speech processor to allow customization for each user was indicated. The programmability also allows the speech processor to be used in comparing the performance of a variety of speech processing algorithms.

It was also found that subject's performance increased as experience with the system was gained. Often many hours of training were involved to reach the reported performance levels. Thus, to be successful the tactile system must be used regularly. Therefore, the prosthesis should eventually be wearable but should definitely be portable during development so that the patients are not limited to "in-lab" use only. This implies that the speech processor must run in real-time and be reasonably compact, with the potential to be wearable. These requirements lead to the use of digital signal processing (DSP) specific microprocessor as the basis of the speech processor hardware.

The problems associated with lipreading and speech production of the deaf, the performance of speech processing algorithms of past systems and the limitations of the tactile sense as a communication channel reported in the literature, led to the conclusion that principal components analysis

(PCA) of speech spectra may be an effective speech processing algorithm for a tactile aid. The principal components algorithm includes both temporal (amplitude envelope) and spectral (frequency tilts) information that have been proven in previous systems (Proctor and Goldstein, 1987; 1985; De Filippo, 1984; Breeuwer and Plomp, 1986; 1984). Moreover, the principal components outputs are well suited to the amplitude modulated, 250 Hz carrier stimulus signal shown to be effective for vibrotactile stimulation (Proctor and Goldstein, 1986). Consequently the principal components analysis of speech spectra was implemented in real-time on the speech processor.

A modular design approach was taken in the speech processor hardware development. The basis of the design is the processor motherboard. This board was designed to be independent of the I/O used in the final system, thereby increasing the utility of the hardware. A number of commercial DSP processors were examined. The TMS320c25 from Texas Instruments was eventually selected due to performance, upgrade path, availability and supporting tools.

After evaluation of the potential resources required by various speech processing algorithms the speech processor mother board was designed. The TMS320c25 was combined with 4k words of external program EPROM and 8k words of data memory, the memory interface logic, crystal oscillator, connector for the application specific I/O daughter board and edge connector for memory expansion allowing up to 36k words of

program and 40k words of data memory. The processor runs at a clock cycle of 100 ns and all memory is accessed with zero wait-states.

The daughter board connector allows full serial and parallel port connectivity to the DSP microprocessor. For the tactile aid implementation the daughter board includes a TLC32044 A/D-D/A circuit, microphone interface, analog output multiplexer and drivers for up to 8 tactile vibrators. Detailed hardware tests were performed using an in-circuit emulator to verify the hardware design of both the speech processor and the I/O circuitry.

In the current implementation, the motherboard measures 13.5 x 10 cm, while the daughter board for tactile I/O is 10 x 6 cm in size. Note that two-sided circuit boards were used to reduce the prototype system cost. At this stage the system is very easily portable. That is the system can go to the users for evaluation, thereby enabling greater training time with the system than if it was tied to a laboratory due to hardware size or non-real time implementations. A wearable implementation using the new TMS320E25, with on-chip program memory is currently underway for the tactile aid. Since the PCA algorithm was carefully implemented to use only on-chip data memory, and the TMS320E25 has a program memory also on-chip, the wearable version of the tactile aid hardware requires only the processor, A/D-D/A chip, and analog output circuitry. It should be noted that the algorithm

implementation on the current system is directly source code compatible with the wearable version.

The real-time program implementation of the PCA algorithm was developed using an IBM-PC compatible and the TMS320c25 Software Development System (plug in card and monitor software). The program consisted of a A/D control, a pre-emphasis filter, 12 channel bandpass filter bank, energy detection for each filter, logarithmic scaling of the energies, principal component weighted summations of the energies, output multiplexer control and generation of the amplitude modulated vibrator signals.

The bandpass filter bank was designed to emulate the characteristics of human audition (Searle, et al., 1979), and resembled the analog filter bank used in the Queen's University tactile vocoder. In addition, the bandpass filters are within 0.5 dB of the ANSI 1/3 Octave filter amplitude specification and thus are similar to those used in many speech research laboratories. Filter implementation structure options were examined to ensure good performance with the fixed-point processor.

The bandpass energy values were approximated by a peak-picking algorithm over a frame of filter output samples. Shorter frame lengths were used for the high frequency group of filters, to retain fast temporal response to the high frequency speech spectra. The logarithm calculation was performed using a 12 segment piecewise linear approximation of the function $20 \cdot \log_{10}(x)$. A program was written to assist

in the determination of the linear segments to within a specified error. The TMS320 log approximation code determines the log of the filter energy with an accuracy of 0.5 dB, as compared to the floating point calculation.

The $1/2$ cosine series approximation to the principal components basis vectors was used in the energy summations (Zahorian, 1988a). The software tasks were distributed over eight sample periods, in order to meet the requirement of real-time execution. The filter code and peak-pick algorithm are executed at the sample rate. The logarithm calculations are executed two per period, and the PC summations are done once per 8 sample periods.

The total program store is 1870 words. The program executes at a sampling rate of 125 μ s, and requires 78% (worst case) of the available real-time period for execution. As the software modules were developed, detailed tests were performed and documented. System level tests demonstrated the PC output signals for various speech inputs. The PC0 output was seen to track the amplitude envelope of the speech signal. The indication of speech spectral tilt was demonstrated by the PC1 response.

5.1 Future Work

Having the electronic modules developed and tested and the principal components analysis algorithm implemented in real-time the logical continuation of the work would involve subject evaluation of the algorithm. Initially tests should

be done with artificially deafened, normally hearing subjects. The evaluations could involve tests both with lipreading and without a view of the speaker. Ideally the tests should be based on speech materials and methods used by other research groups in order that a performance comparison with other systems be possible.

Evaluation methods and materials developed at the University of Iowa (Tyler et al., 1986) seem to be well suited. Tyler's group has performed a number of evaluations of cochlear implant systems with a variety of patients and have made their speech test materials available to other research groups. The training methods and materials employed by Queen's University in the evaluation of their tactile vocoder (Brooks, 1984) should also be examined as candidates for testing the PCA tactile aid. If after evaluation with artificially deafened subjects the algorithm shows promise, tests could begin with deaf subjects. The Audiology department at the Children's Hospital of Eastern Ontario (CHEO), in Ottawa, has expressed an interest (Smith, 1988), and given their expertise with profoundly deaf children, would be an ideal location for future evaluations.

In addition to evaluation of the PCA algorithm as implemented there are a number of variables in the implementation that could be investigated. The first of these is the frequency range of analysis. The current system is bandlimited to 4 kHz. Although this is sufficient for telephone quality speech, an additional 2-4 kHz at the high

frequency end may provide information about higher frequency consonant sounds.

Variations in the tactile display and the area of tactile perception using the vl220 vibrators should be examined in greater detail than the scope of this thesis allowed. Formal studies of tactile amplitude and vibrator spatial location perception may provide the information necessary to develop an optimum output scheme for the PCA algorithm. The issue of whether to code the PC positive and negative values via two transducers or via a single transducer with the PC value dc offset should be addressed. In the two transducer coding scheme, variations in transducer placement should also be tested.

The effects of various automatic gain control schemes and noise reduction methods should be examined. In a real-world environment (i.e. not a quiet classroom or laboratory), some type of noise reduction may be required. The effects of noise on the PCA algorithm have not been investigated but should be understood in order to enhance the performance of the system in noisy conditions.

On the hardware side of the project, a reduction in the size of the system is currently underway. To meet the objective of a wearable electronic module the speech processor will use the Eprom based TMS320E25 DSP micro-processor. Additional work towards reducing the part count of the analog I/O circuitry will also need to be done. Power consumption reduction techniques, such as power-down modes,

reduced power supply voltage, reduced clock speed, improved transducers and stimulation schemes should be investigated in order to reduce the size of the required power supply in the wearable version.

5.2 Concluding Remarks

In conclusion, a digital signal processing based electronic speech processor hardware has been developed for use in auditory prosthetic systems. The speech processor is compact and powerful enough to implement very sophisticated signal processing algorithms in real-time. The modular design of the hardware allows different input/output circuitry to be easily connected to the speech processor.

A tactile I/O module was developed to interface to the speech processor. This module contains a 14 bit A/D-D/A converter and hardware to drive up to 8 vibrotactile transducers. The programmability of the combined system will allow various signal processing algorithms to be implemented in real-time and evaluated for use in a tactile aid.

A detailed review of the problems of profound deafness and previous work in the area of tactile aids suggested that the principal spectral components analysis of speech spectra may be well suited to the needs of the profoundly deaf patient. This algorithm was implemented on the speech processor hardware and tactile I/O module. The PCA algorithm consisted of a bandpass filter bank based on human audition, filter energy estimation, logarithmic scaling and principal component summation. In addition, software was developed to control the I/O and generate the tactile stimulus signals. Responses of the system to speech input demonstrated the potential of principal components analysis for a tactile prosthesis.

Appendix I **Phoneme Symbols and Word Context**

Consonants:

Phoneme	Word
/b/	be
/tʃ/	chair
/k/	chemist, key
/d/	dog
/f/	full
/g/	got
/dʒ/	gym, jem
/ŋ/	ring
/ʒ/	rouge
/h/	help
/l/	lamp
/m/	mat
/n/	nap
/p/	pet
/r/	rope
/s/	soap
/ʃ/	shoe
/t/	two
/θ/	three
/ð/	these
/v/	vary
/w/	was
/hw/	why
/j/	yes
/z/	zoo

Vowels:

Phoneme	Word
/ɑ/	father, pot
/e/	age, bait
/ɛ/	air, bet
/æ/	add
/i/	even
/I/	bit
/ɜ/	bird
/aI/	ice
/o/	moat
/ʊ/	pun
/ʌ/	but
/aʊ/	pout
/ɔI/	oil

Appendix II-A
TMS320c25 - Memory Interface Timing

Program Memory

The program memory decode circuit and timing diagram may be found in Figure 46, in this Appendix. The design calculations are shown below.

Parameters: 74AS27 - $t_{PLH} = 5.5 \text{ ns max}$
 $t_{PHL} = 4.5 \text{ ns max}$

CY7C291-25 - $t_{ACS} = 20 \text{ ns max}$
 $t_{AA} = 25 \text{ ns max}$ all into a 5 pF
 $t_{HZCS} = 20 \text{ ns max}$ load

Data line load:

2x CY7C291-25 + 1x P4C164 + I/O con. + PCB tracks
= 2(8) + 7 + ~5 + ~5
= 33 pF

Modified timing for CY7C291-25 - $t_{ACS} = 21 \text{ ns max}$
 $t_{AA} = 26 \text{ ns max}$
 $t_{HZCS} = 21 \text{ ns max}$

Address line load:

4x CY7C291-25 + 2x P4C164 + 4 TTL gates + I/O con. + PCB
= 4(8) + 2(7) + 4(10) + ~5 + ~7
= 98 pF

The TMS320c25 timing is specified for a 100 pF load, therefore no correction is necessary.

Read Access Time:

Time to CS1* from c25 address valid = $t_{SU}(A) = 13 \text{ ns}$.
Time to CS1* from c25 STRB* valid = 0 ns.
Time to CS2 from c25 address valid = $t_{PLH} = 5.5 \text{ ns}$.
Time to CS3 from c25 address valid = 0 ns.

∴ CS1* is the limiting factor.

Figure 46 Program Memory Decode Circuit and Timing

CY7C291-25 output worst case data valid from c25 address valid:

$$t_{SU}(A) + t_{ACS} = 13 + 21 = 34 \text{ ns max.}$$

The TMS320c25 requires data $t_a(A) = 40 \text{ ns max}$ from address valid.

∴ Timing margin is $40 - 34 = 6 \text{ ns min.}$, and read timing specifications are met with respect to EPROM chip select valid timing.

Check data valid from c25 STRB* timing:

CY7C291-25 supplies data $t_{ACS} = 21$ from CS1*, which is driven by STRB*.

The TMS320c25 requires data $t_a(A) - t_{SU}(A) = 40 - 13 = 27 \text{ ns max}$, from STRB* valid.

∴ Timing margin is $27 - 21 = 6 \text{ ns min.}$, and read timing specifications are met with respect to TMS320c25 STRB* valid timing.

Check data valid from CY7C291-25 address valid:

CY7C291-25 provides data $t_{AA} = 26 \text{ ns max}$, from address valid.

The TMS320c25 requires data $t_a(A) = 40 \text{ ns max}$ from address valid.

∴ Timing margin is $40 - 26 = 14 \text{ ns min.}$, and read timing specifications are met with respect to EPROM address valid timing.

Check for Bus Contention:

Read from EPROM followed by a write:

CY7C291-25 tristates the data bus $t_{HZCS} = 21 \text{ ns max}$, after a chip select (CS) is disabled. Thus the EPROM remains on the bus 21 ns into the following cycle, since CS1*, driven by STRB* is disabled at the end of the read cycle.

TMS320c25 can drive the bus $t_{en}(D) = 0 \text{ ns min}$, after $STRB^*$ goes low, but the $STRB^*$ signal stays high for 50 ns min, between cycles.

$\therefore$ No bus contention.

Write to data memory, followed by a Read from EPROM:

TMS320c25 drives the data bus $t_{dis}(D) = 40 \text{ ns}$ after $STRB^*$ is high, during a memory write cycle.

Again, $STRB^*$ remains high for 50 ns between cycles, so the TMS320c25 will be off the data bus, $50 - 40 = 10 \text{ ns}$, before the EPROM tries to drive the bus.

$\therefore$ No bus contention.

Data Memory

The data memory decode circuit and read timing diagram may be found in Figure 47. The write cycle timing is shown in Figure 48. The design calculations are shown below.

Parameters:

74AS27	- $t_{PLH} = 5.5 \text{ ns max}$	
	$t_{PHL} = 4.5 \text{ ns max}$	
74AS00	- $t_{PLH} = 4.5 \text{ ns max}$	
	$t_{PHL} = 4.0 \text{ ns max}$	
P4C164-25	- $t_{AC} = 25 \text{ ns max}$	
	$t_{AA} = 25 \text{ ns max}$	all into a 30 pF
	$t_{OHZ} = 12 \text{ ns max}$	load
	$t_{OE} = 13 \text{ ns max}$	

Timing correction for capacitive loading is not required here, since the specifications are provided for the actual circuit load.

Read Access Time:

Time to $CE1^*$ from c25 address valid = $t_{PLH'27} + t_{PHL'00}$
 $= 5.5 + 4.0$
 $= 9.5 \text{ ns max}$

Time to OE^* from c25 address valid = $t_{SU}(A) = 13 \text{ ns max}$

Data Memory Decode

Figure 47 Data Memory Decode Circuit and Read Timing

3 of/de 3

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS
STANDARD REFERENCE MATERIAL 1010a
(ANSI and ISO TEST CHART No. 2)

P4C164-25 output data valid from c25 address valid, and CE1* valid:

$$t_{PLH'27} + t_{PHL'00} + t_{AC} = 9.5 + 25 = 34.5 \text{ ns max.}$$

The TMS320c25 requires data $t_a(A) = 40 \text{ ns max}$ from address valid.

∴ Timing margin is $40 - 34.5 = 5.5 \text{ ns min.}$, and read timing specifications are met with respect to RAM chip select valid timing.

Check data valid from c25 STRB* timing:

P4C164-25 supplies data $t_{OE} = 13$ from OE*, which is driven by STRB*.

The TMS320c25 requires data $t_a(A) - t_{SU}(A) = 40 - 13 = 27 \text{ ns max}$, from STRB* valid.

∴ Timing margin is $27 - 13 = 14 \text{ ns min.}$, and read timing specifications are met with respect to TMS320c25 STRB* valid timing.

Check data valid from P4C164-25 address valid:

P4C164-25 provides data $t_{AA} = 25 \text{ ns max}$, from address valid.

The TMS320c25 requires data $t_a(A) = 40 \text{ ns max}$ from address valid.

∴ Timing margin is $40 - 25 = 15 \text{ ns min.}$, and read timing specifications are met with respect to RAM address valid timing.

Check for Bus Contention:

Read from RAM followed by a write:

P4C164-25 tristates the data bus $t_{OHZ} = 12 \text{ ns max}$, after STRB* is disabled.

TMS320c25 can drive the bus $t_{en}(D) = 0 \text{ ns min}$, after STRB* goes low, but the STRB* signal stays high for 50 ns min, between cycles.

∴ No bus contention.

Write Cycle Timing:

Parameters: P4C164-25 - t_{CW} = 18 ns min
 t_{AW} = 18 ns min
 t_{WP} = 18 ns min
 t_{DW} = 13 ns min
 t_{DH} = 0 ns min
 t_{WZ} = 10 ns max
 t_{OW} = 3 ns min

Write Access Time:

Time to CE1* from c25 address valid = $t_{PLH'27} + t_{PHL'00}$
= 5.5 + 4.0
= 9.5 ns max.

Time to WE* from c25 R/W* valid = 0 ns max.

TMS320c25 provides data $t_h(A) + t_{SU}(D)W$ before R/W* invalid
= 17 + 30
= 47 ns.

P4C164-25 requires data $t_{DW} = 13$ ns min, before R/W* invalid.

∴ P4C164-25 data setup time requirements are met with a margin of $47 - 13 = 34$ ns.

Check data hold time after WE* invalid:

P4C164-25 requires data be held $t_{DH} = 0$ ns min, after the WE* signal is high.

∴ TMS320c25 does not need to hold the data longer than the end of the R/W* signal valid.

Figure 48 Data Memory Write Timing

Check for Bus Contention:

Note: the OE will be enabled during the write cycle, since it is connected to STRB*. This could cause problems, but the WE* signal overrides the OE* signal and tristates the data outputs. Providing the outputs are in tristate during the time the c25 drives the data bus the design will function correctly.

Check for STRB* (OE*) signal problems at beginning of write cycle:

As shown by the RAM Data Out signal in Figure 48, the RAM data bus is forced into high impedance mode, $t_{WZ} = 10 \text{ ns max}$, after the R/W* signal goes low.

The TMS320c25 STRB* signal, which will try to enable the RAM data bus via OE*, is valid $t_{SU}(A) = 13 \text{ ns}$, after the R/W* signal goes low.

The TMS320c25 begins driving the data bus,

$$t_{SU}(A) + t_{en}(D)$$

$$= 13 + 0$$

= 13 ns min after the R/W* signal goes low (and address valid).

Margin $\rightarrow 13 - 10 = 3 \text{ ns}$ between RAM data bus changing to high impedance and the TMS320c25 writing on the data bus.

$\therefore$ the R/W* signal will override the OE* signal (driven by STRB*) and force the RAM data bus to be in high impedance during beginning of the TMS320c25 write cycle.

Check for STRB* (OE*) signal problems at beginning of write cycle:

As shown by the RAM Data Out signal in Figure 48, the RAM data bus will exit from high impedance mode, $t_{OW} = 3 \text{ ns min}$, after the R/W* signal goes high, at the end of the write cycle.

The TMS320c25 STRB* signal, which enables the RAM data bus via OE*, changes to invalid $t_h(A) = 17 \text{ ns min}$, before the R/W* signal goes high. And the RAM outputs are forced to high impedance, $t_{OHZ} = 12 \text{ ns max}$ after STRB* is high.

Thus the RAM data bus is forced to high impedance,

$$t_h(A) - t_{OHZ} = 17 - 12 = 5 \text{ ns min, before the R/W* signal goes high.}$$

$\therefore$ the RAM data bus will be forced to high impedance mode by the OE* signal invalid, $5 + 3 = 8 \text{ ns min}$, before the RAM data bus could potentially be out of high impedance mode due to the R/W* (WE*) signal going high at the end of the write cycle.

Check for memory write followed by read contention:

Write to data memory, followed by a Read from EPROM.

TMS320c25 drives the data bus $t_{dis}(D) = 40$ ns after $STRB^*$ is high, during a memory write cycle.

Again, $STRB^*$ remains high for 50 ns between cycles, so the TMS320c25 will be off the data bus, $50 - 40 = 10$ ns, before the EPROM tries to drive the bus.

∴ No bus contention.

Ready Signal Generation

As shown in Figures 46-48, the TMS320c25 requires a READY signal input during the memory accesses. Figure 49 shows the circuitry used to generate the necessary signal.

Figure 49 READY Signal Generation Circuit

Time from address valid to READY valid required by the TMS320c25:

$$\begin{aligned}
 &= t_{su}(A) + t_d(SL-R) \\
 &= 13 + 5 \\
 &= 18 \text{ ns min.}
 \end{aligned}$$

The READY generation circuit provides the READY signal in

$$= t_{PLH00c} + t_{PHL00b} + t_{PLH00d}$$

$$= 4.5 + 4 + 4.5$$

$$= 13 \text{ ns max}$$

after the address is valid.

$\therefore$ The timing requirements of the TMS320c25 are met.

Appendix II-B
DSP Motherboard Hardware Schematic Diagram

The schematic diagram of the speech processor module is shown in Figure 50.

Figure 50 Speech Processor Board Schematics - CPU

Figure 50 Speech Processor Board Schematics - Program Memory

Figure 50 Speech Processor Board Schematics - Data Memory

Appendix II-C
DSP Motherboard PCB Layout

The printed circuit board (PCB) layout of the speech processor module is shown in Figure 51.

Figure 51 Speech Processor PCB Layout

Appendix II-D

DSP Board Test Programs and Results

The first DSP board tests involved using the TI XDS in - circuit emulator memory FILL and display DDM commands to test memory access. In addition, four test programs were written to verify the proper functioning of the speech processor hardware design and PCB manufacture. The programs were run via the in-circuit emulator system, with code downloaded from the IMB-PC compatible computer. The programs read and write from data memory and read from program memory. The code listings and logic analyzer photographs of appropriate signals follow.

The first test program, dmtest.asm listed below, loads the data memory locations with all ones and reads back each location to verify the write. If an error is encountered, the location of the error and the memory content at that location are saved. Variations of this test were also performed with different bit patterns. Figure 52 is a photograph of an Arium logic analyzer display taken during the execution of the dmtest.asm program. A write cycle followed by the read cycle can be seen in the figure.

```

.....
*  HARDWARE DATA MEMORY TEST
.....
*dmtest.asm
.
    AORG    0
    B      START
.
    AORG   >32
START
    LDPK 4
    LARP 0
    LRLK 0,>200          *LOAD AR0 WITH START OF B0
    SOVM
    RSXM
    ZAC
    RPTK 128
    SACL *+              * CLEAR BLOCK B0
.
    LRLK 0,>4000          * LOAD AR0 WITH TEST MEM. START
    LRLK 1,>200           * LOAD AR1 WITH B0 START
    LRLK 2,>2000          * AR2 COUNTS DOWN MEMORY ACCESSSES
.
* START OF MEMORY ACCESS LOOP
.
LOOP    LALK >FFFF       * LOAD LO ACCUM. WITH FFFF
.

```

```

      SACL *          * STORE LOW ACCUM. INTO DMEM
      AND *          * AND STORED VALUE WITH ACCUM.
      * IF ARE THE SAME - ACC. = 0000FFFF
      * IF ARE DIFF. - ACC. = 0000xxxx
      SBLK >FFFF      * IF MEM OK THEN ACC. = 0
      BNZ BAD          * IF ACC. NOT EQ. 0 GO TO BAD
      ADRK >1          * INC AR0
      LARP 2
      BANZ LOOP,*-,0   * CHECK IF ALL MEM. LOC. TESTED
DONE  IDLE             * IF DONE THEN IDLE
      B DONE
      AORG >50
      BAD LARP 1
      SAR AR0,*+,0     * STORE ADDR. OF BAD MEM., INC AR1
      LAR AR3,*+,1     * PUT BAD MEM. CONTENTS IN AR3, INC AR0
      SAR AR3,*+,2     * STORE BAD MEM. CONT. IN B0, INC AR1
      BANZ LOOP,*-,0   * CHECK IF DONE
DN1   IDLE
      B DN1
      END

```

Figure 52 Logic Analyzer Display of Write Followed by Read

The minimum resolution of the logic analyzer was 10 ns. In order to obtain a more accurate measure of the program memory decode delay, the test program below was created. This program cycles continuously through a loop performing two writes to the data memory. A Tektronix 475 oscilloscope was used to monitor the DS* signal from the TMS320c25 and the data memory chip select, CE1* at the P4C169. Figure 53 is a photograph taken from the scope display. Inverting probes were used in this case so the signals are shown. The DS* signal is the trace that rises first, followed by the CE1* signal, after the propagation delay through the decode logic. The times scale of the photograph is 5 ns per division, and it can be seen that the address decode delay is approximately 7 ns. Recall that the design worst case time was 9.5 ns maximum.

.....

* chip select delay time test program

```

.      AORG >0
      B      START
.
      AORG >32
.
START  LARP      0
      LRLK AR0,>4000
      LRLK AR1,>8200
      RSXM
      LALK >FFFF,0
LOOP   SACL *,0,1
      SACL *,0,0
      B      LOOP
      END

```

Figure 53 Data Memory Address Decode Propagation Delay

The program memory access test program, `pro.asm` is listed below. `pro.asm`, performs a continuous loop read from program memory. Figure 54 shows a logic analyzer photograph of the program memory enable signal for the 0-2k bank of memory. Figure 55 is an oscilloscope display showing the program memory decode circuit delay. In this figure, the PS^* signal is changing from high to low, and approximately 4 ns later (2ns per division), the $CS2$ signal rises. Recall that the worst case decode delay was 5.5 ns maximum.

```
*pro.asm
    AORG >0
    B    START
.
    AORG >32
START LDPK 4
      LARP 0
      SOVM
      RSXM
LOOP  BLKP >100,>0
      SBLK >FFFF
      B    LOOP
      END
```

Figure 54 Program Memory Access Timing from Logic Analyzer

Figure 55 Program Memory Address Decode Propagation Delay

Appendix II-E TMS320c25 Program to Test the AIC Interface

The AIC TLC32045 chip interface and setup software was tested using a special testbed connected to the XDS. The testbed hardware provided the TLC32045 with a voltage reference, I/O signals to BNC jacks, a TMS320c25 socket, and power supply conditioning. This test program was later run on the target analog daughter board to verify the interface and proper functioning of the TLC32045. The program is a modified version of a listing provided in the TLC3204x user guide (Texas Instruments, 1988) and is shown below:

```

.....
*   TLC32040 COMMUNICATION PROGRAM
*
*   from TLC32040 manual page B-4
*   Modified by : Tim Rahrer
*
* 88 09 16                      V2.0
.....

* Memory mapped register:

DRR EQU 0      * data receive register address (pg 4)
DXR EQU 1      * data transmit " "
IMR EQU 4      * interrupt mask register " "

SSTST EQU 96   * store ST0 on page 0
INTST EQU 97   * " ST1 " "

* Reserved on-chip RAM as flags and storages
* ( block B0 511 thru 530 )

FXE EQU 5      * ignore first XINT flag
FRE EQU 6      * ignore first RINT flag
TMP0 EQU 7     * temporary register
ACCHST EQU 8   * stack for ACCH
ACCLST EQU 9   * " " ACCL
NSAMP EQU 10   * number of samples to be aquired
STMEM EQU 11   * starting data mem. addr. for storage
RVECT EQU 12   * vector for RINT
XVECT EQU 13   * " " XINT
VRCV EQU 14    * RINT vector storage
VNRM EQU 15    * XINT " "
VS1 EQU 16     * secondary vector storage1
VS2 EQU 17     * " " storage2
DRCV EQU 18    * receive data storage
DXMT EQU 19    * transmit " "
D2ND EQU 20    * secondary " "
FRCV EQU 21    * receive flag

```

```

FXMT EQU 22      * transmit *
F2ND EQU 23      * secondary comm. flag
TA    EQU 24      * TA counter value for AIC
TB    EQU 25      * TB " " " "
CTRL EQU 26      * control word for AIC sampling freq. etc.
.
      AORG 0
      B    START      * jump to initialization routine
.
.....RECEIVE INTERRUPT LOCATION .....
.
      AORG 26
      B    RINT      * jump to receive interrupt routine
.
..... TRANSMIT INTERRUPT LOCATION .....
.
      AORG 28
      B    XINT      * jump to transmit interrupt routine
.
..... START INITIAL PROGRAM .....
.
      AORG >32
.
START DINT
      CALL INIT      * call initializing routine
      CALL IGR      * call routine to ignore the 1st two inter.
.
      LDPK 4
      LARK 7,2      * AR7 contains loop repeat number
      LRLK 6,>218    * AR6 points to TA,TB then CTRL
      LARP 6
.
LOOP  LAC  *+,0,7      * load first AIC ctrl. word into acc.,
      * decr. AR6, point to AR7
      SACL D2ND,0      * store 2nd data in D2ND
      LALK >FFFF      * load 11 into d0,d1 to prep. for 2nd. comm.
      SACL DXMT,0      * store in DXMT for use by C2ND
      LDPK 0
      SACL DXR,0      * store 2nd comm. code in transmit stor. reg.
      CALL C2ND      * check for secondary comm. code, jump to S1
      * if code d0,d1 = 1,1
.
      LDPK 4
WAIT  LAC  F2ND,0      * load 2nd comm. flag
      ANDK >FFFF,0      * AND flag with ones, if waiting, F2ND
      * will be ones and acc. = 0
.
      BNZ WAIT
      ZAC
      SACL FXMT      * clear transmit flag
      SACL F2ND      * clear 2nd transmit flag
      BANZ LOOP,*-,6  * if have not sent all 3 ctrl words >
      * loop thru again, decr. AR7, point to AR6
.

```

```

* .....
*
* user area      set FXMT and FRCV to 0
*                count and store samples
*                sin x/x
* .....
*
*      ZAC
*      LDPK 4
*      LAR 2,NSAMP      * load AR2 with the number of samples
*      LAR 3,STMEM      * " AR3 " " starting addr.
MLOOP  LAC FRCV,0      * load the receive flag
*      ANDK >FFFF,0    * check if flag set, if not
*      BZ MLOOP        * then loop
*      LARP 3          * ELSE
*      SSXM
*      ZAC
*      LAC DRCV,0      * get new input data sample
*      SACL *,0,2      * store in dmem ptd to by AR3, decr AR3
*                      * point to AR2
*
*      ZAC
*      SACL FRCV,0      * clear receive flag
*
*      TEST LAC FXMT,0  * load transmit flag
*      ANDK >FFFF,0    * check if flag is set
*      BZ TEST          * if not set then loop to test
*      ZAC
*      LAC DRCV,0      * if set > get last input sample
*
* can implement ( sin x ) / x here
*
*      LDPK 0
*      SACL DXR,0      * put input sample into output register
*      ZAC
*      LDPK 4
*      SACL FXMT,0      * clear transmit flag
*
*      BANZ MLOOP,*,3  * if number of samples to be collected not
*                      * = 0 then loop back for more, decr AR2
*
*      RXF              * set xf flag to signal end
DONE  IDLE
      B      DONE

```

```

* .....
*
*      INITIALIZING ROUTINE
*
* This routine initializes the status regs., flags,
* vector storage contents and internal data locs.
* 511 to 529.
*

```

```

.....
.
INIT  LDPK  4          * dp = 4
      LALK  >0E04,0    * 0000 1110 0000 0100B
      SACL  TMP0,0
      LST   TMP0       * ov = 0 clear overflow
                      * ovm = 1 set overflow mode
                      * ? = 1 not affected
                      * intm = 1 not affected
                      * dp = 0000 0100 = 4
                      * lst loads status reg.

      LALK  >03F0      * 0000 0101 1111 0000B
      SACL  TMP0,0
      LST1  TMP0       * arb = 0
                      * cnf = 0 ( B0 is data mem. )
                      * tc = 0 ( test/control flag bit)
                      * sxm = 1 enable sign ext. mode
                      * c = 0 carry set to 0, fsm = 1
                      * xf = 1
                      * f0 = 0 16 bit transfer mode
                      * txm = 0 (fsx is input)

      LDPK  0          * DP = 0
      ZAC                      * clear registers
      SACL  DXR,0
      SACL  DRR,0
      LDPK  4          * DP = 4
      LRLK  AR0,>205
      LARP  0          * clear block B0
      RPTK  22         * repeat 22 times starting from 5 on pg 4
      SACL  '+,0

      LALK  >1020      * load xx01 000x x010 0000 to set counter A
      SACL  TA,0        * to SCF freq= 312500 Hz ( 288 kHz )
      LALK  >4E9E      * load x100 111x 1001 1110 to set counter B
      SACL  TB,0        * to sample freq = 8013 Hz
      LALK  >00A7      * load xxxx xxxx 1010 0111 to set AIC ctrl. wd.
      SACL  CTRL,0     * for gain = 2, synch mode, BPF in, +- 3 V in

      LALK  >3FFF      * starting mem. addr. for data storage is
      SACL  STMEM,0    * 3FFF
      LALK  1024      * load number of samples to collect
      SACL  NSAMP,0

.
.
.
Interrupt Masking
.
      LDPK  0
      LACK  >30        * 0000 0000 0011 0000B
      SACL  IMR,0      * XINT-----|| |||
                      * RINT-----| |||
                      * TINT-----|||
                      * INT2-----|||

```

```

*
*
*      INT1-----||
*      INT0-----|
*
LDPK  4
LALK  NRM,0      * normal xint routine address
SACL  VNRM,0     * store in VNRM
*
LALK  S1,0       * secondary xint address 1
SACL  VS1,0     * store in VS1
*
LALK  S2,0       * secondary xint address 2
SACL  VS2,0     * store in VS2
*
LALK  RCV,0      * rint routine address
SACL  VRCV,0    * store in VRCV
*
LALK  IGNRR,0    * set ignore first rint address
SACL  RVECT,0   * store in RVECT
*
LALK  IGNRX,0    * set ignore first xint address
SACL  XVECT,0   * store in XVECT
RET
*
* .....
*
* RECEIVE INTERRUPT ROUTINE
*
* This routine stores receive data in loc.
* *DRCV ( 524 on pg 4 ) and sets rec. flag FRCV*
* Two levels of nesting are allowed without
* * stack extension.
* .....
*
RINT  LDPK  0
      SST  SSTST      * push ST0 register
      DINT
      LDPK  4          * DP = 4
      SACL  ACCLST,0   * push ACCL
      SACH  ACCHST,0   * " ACCH
      LAC   RVECT,0    * load vector address into acc. (RCV)
      CALA                      * store current pc, load ACCL into pc
      ZALS  ACCLST     * pop accum.
      ADDH  ACCHST
      LDPK  0
      LST   SSTST      * pop ST0
      LDPK  4
      EINT                      * enable interrupts
      RET
*
RCV   LDPK  0
      LAC   DRR,0       * load input data from Data Receive Reg.
      LDPK  4
      SACL  DRCV,0     * store input data in DRCV

```

```

LACK >FF      * set receive flag
SACL FRCV
RET

```

``` XMIT INTERRUPT ROUTINE ```

```

* This routine writes xmit data ( the cont.
* of DXMT ) to the DXR reg. according to com.
* condition, i.e. normal communication of
* secondary. For normal comm. call NRM rout.
* For secondary comm. call routines S1 and S2
* Again two levels of nesting are allowed
* without stack ext.

```

```

XINT  LDPK 0
      SST  SSTST      * push ST0
      DINT
      LDPK 4          * dp = 4
      SACL ACCLST,0    * push ACCL
      SACH ACCHST,0    * " ACCH
      LAC  XVECT,0     * load xmit vector address
      CALA          * jump to transmit routine
      ZALS ACCLST      * pop accum.
      ADDH ACCHST
      LDPK 0
      LST  SSTST      * pop ST0
      LDPK 4
      EINT          * enable interrupt
      RET

```

``` NORMAL DATA WRITING ROUTINE ```

```

* This routine is called when normal comm.
* occurs. Xmit data is written to DXR and
* XMIT flag is set.

```

```

NRM   LAC  DXMT,0      * move data to be xmitted from DXMT register
      LDPK 0
      SACL DXR,0       * to Data Xmit Register
      LDPK 4
      LACK >FF        * set Xmit flag
      SACL FXMT
      RET

```

``` SECONDARY DATA WRITING ROUTINE 1 ```

- * This routine is called when a secondary
- * comm. occurs. Secondary data is written
- * to DXR and content of XVECT is modified
- * for continuing the secondary communication *

```

.....
*
S1  LAC    D2ND,0          * move secondary date to xmit from D2ND
    LDPK   0
    SACL   DXR,0          * to DXR
    LDPK   4
    LAC    VS2,0          * load transmit vector address to
    SACL   XVECT,0        * go to secondary writing routine 2
    RET

```

```

.....

```

SECONDARY DATA WRITING ROUTINE 2

- * This routine is called when second. comm.
- * occurs. This routine writes dummy data to
- * DXR to clear the secondary code for the
- * protectin of double writing the sec. code
- * It resets secondary flag (F2ND) and
- * modifies the content of XVECT for normal
- * communication.

```

.....
*
S2  ZAC                      * clear DXR for protection from
    LDPK   0
    SACL   DXR,0          * double secondary commun.
    LDPK   4
    SACL   F2ND          * clear secondary flag
    LACK   >FF          * set
    SACL   FXMT,0        * xmit end flag
    LAC    VNRM,0        * set normal communication transmit vector
    SACL   XVECT,0
    LDPK   0
    LAC    INTST,0       * enable all interrupts
    SACL   IMR,0
    LDPK   4
    RET

```

```

.....

```

CHECK SECONDARY CODE

- * This routine checks whether the data in
- * DXMT has secon. code or not. If sec. code
- * exists the disable maskable interrupts
- * except XINT, modify the contents of XVECT
- * for sec. comm. and set F2ND. Recommended
- * for sending control words to AIC

```

.....

```



```

C2ND  LDPK  4          * dp = 4
      LACK  03          * get comparing mask
      SACL  TMP0        * save in TMP0
      LAC   DXMT,0      * load transmit data to check for sec. comm.
      AND   TMP0        * check if d0,d1 = 1,1 for secondary comm.
      SUB   TMP0,0      *
      BZ    C2ND1       * if yes then go to C2ND1
      RET

```

```

C2ND1  LACK  >FF        * set secondary flag
      SACL  F2ND,0      *
      LACK  >20         * enable only XINT
      LDPK  0
      SACL  IMR,0
      LDPK  4
      LAC   VS1,0       * modify vector address for secnd. comm.
      SACL  XVECT,0
      LAC   DXMT,0      * load primary data in DXMT
      LDPK  0
      SACL  DXR,0       * send current primary data before secondary
      LDPK  4
      RET

```

```

.....

```

```

* CHECK FIRST INTERRUPT *

```

```

* This routine checks whether both first *
* interrupts have occurred. If this routine *
* is called after reset, this routine waits *
* for both interrupts and then returns *
.....

```

```

IGR  LARK  0,10
     LARP  0
     FXF
RESL SBRK  1          * set xf pin low to reset AIC
     BANZ  RESL,*     * dec. AR0 by 1
     SXF
     EINT
IGRL LAC   FRE,0      * check first interrupt after
     BZ    IGRL       * master reset
     LAC   FXE,0
     BZ    IGRL
     RET

```

```

.....

```

```

* IGNORE INTERRUPT ROUTINE *

```

```

* These routines ignore the first RINT and *
* XINT after DSP reset. The routines only *
* set flags and modify each vector addr. to *
* normal interrupt addr. but do not read or *

```

* write to serial ports.
 * NOTE: THE FIRST DATA THAT THE AIC WILL
 * RECEIVE AFTER DSP RESET IS 0000H

*
 IGNRR LACK >FF
 SACL FRE,0
 LAC VRCV,0 * set normal receive return address
 SACL RVECT,0
 RET

*
 IGNRX LACK >FF
 SACL FXE,0
 LAC VNRM,0 * set normal transmit address
 SACL XVECT,0
 RET

*
 END

Appendix II-F

Transducer Driver Schematics and Test Information

The transducer driver circuit schematic is shown in Figure 57. Initial testing was done using XDS commands to output a control word to change the state of the analog switches. A program, ANA14R.ASM, listed below, alternated between enabling SW1, while disabling SW4 and the opposite. The logic analyzer photograph in Figure 56 shows the input STRB*, IS*, XF, data lines D10 and D6 and the CLK enable to the 74HCT534. It can be seen that the turn on time of the DG201 (SW4 low-high transition in Figure 56) is faster than the turn off time (SW1 high to low transition).

```

.      ..... ANA14R.ASM .....
.
.  Test program for interfacing c25 to analog switches
.
      AORG >0
        B      START
        AORG >50
.
START  RXF
      LARP 0
      LRLK 0,>200      * LOAD ADDRESS OF CTRL WORD INTO AR0
      LALK >0040      * LOAD CTRL WORD INTO ACCUM.
.
.
.      * >0400 SWITCH 1 IS ON
.      * >1000 " 2 "
.      * n/c " 3 "
.      * >0040 " 4 "
.
      SACL *+,0
      LALK >0400
      SACL *-,0
LOOP  OUT *+,0,1

      LRLK 1,>15      * AR1 USED FOR DELAY

WAIT  LARP 1
      BANZ WAIT,*-,0

      OUT *-,0,1

      LRLK 1,>15      * AR1 USED FOR DELAY

WAIT  LARP 1
      BANZ WAIT,*-,0

      B      LOOP
      END

```

Figure 56 Analog Driver Circuit 1/2

Figure 57 Analog Driver Circuit 2/2

Appendix III Vibrator Response Curves

The frequency response for four V1220 transducers are shown in Figures 58-61. The measurement technique is outlined in Chapter 4.

Figure 58 Frequency Response of Transducer #1

Figure 59 Frequency Response of Transducer #2

Figure 60 Frequency Response of Transducer #3

Figure 61 Frequency Response of Transducer #4

Appendix IV-A **Fortran Filter Design Program**

```

C-----
C   THIS IS A IIR FILTER DESIGN PROGRAM
C   TAKEN FROM PAGE 319 OF DIGITAL FILTER DESIGN
C   BY PARKS AND BURRUS. THE PROGRAM WAS WRITTEN
C   BY C.S. BURRUS, RICE UNIVERSITY, JAN 1987
C
C   Modifications made by: Tim Rahrer, 1988
C
C-----
      DIMENSION PR(20), PI(20), ZR(20),ZI(20)
      DIMENSION B1(20), B2(20), A1(20), A2(20)
      DIMENSION FM(530)
      COMMON /PARM/R1,R2,WP,WS,N2,N,SR,KA,KAD,KOD,KF
      COMMON /ROOT/PR,PI,ZR,ZI

C-----INPUT SPECS, PREWARPING AND PREFREQXFRMING-----
      PRINT *, 'ENTER NUMBER OF FREQS TO DISPLAY '
      READ *, KK
      PRINT *, ' '
10    PRINT *, 'ENTER 1 FOR BW, 2 FOR CHEBY, 3 FOR ICHEBY, 4 FOR ELL '
      READ *, KA
      PRINT *, ' '
      PRINT *, 'ENTER 1 FOR LOWPASS, 2 FOR HP, 3 FOR BP, 4 FOR BR '
      READ *, KF
      PRINT *, ' '
      PRINT *, 'ENTER 1 FOR ANALOG, 2 FOR DIGITAL '
      READ *, KAD
      TP = 6.283185307179586
      IF (KAD.EQ.1) GOTO 12
      PRINT *, ' '
      PRINT *, 'ENTER THE SAMPLE RATE '
      READ *, SR
      PRINT *, ' '
12    IF (KA.EQ.4) GOTO 20
13    PRINT *, 'ENTER THE ORDER '
      READ *, N
      PRINT *, ' '
      IF (KF.GE.3) GOTO 15
      PRINT *, 'ENTER THE BAND EDGE IN UN-NORMALIZED HZ '
      READ *, FP
      PRINT *, ' '
      WP = PREWRP(TP*FP)
      IF (KF.EQ.2) WP = 1.0/WP
      IF (KA.EQ.1) GOTO 30
      PRINT *, 'ENTER PASSBAND RIPPLE OR STOPBAND ATT IN POSITIVE DB '
      READ *, R1
      PRINT *, ' '
      IF (KA.EQ.3) WS = WP
      GOTO 30
15    PRINT *, 'ENTER THE LOWER & UPPER BAND EDGES IN HZ '

```



```

      READ *, F1, F2
      PRINT ' '
      W1 = PREWRP(TP*F1)
      W2 = PREWRP(TP*F2)
      W0 = SQRT(W1*W2)
      WP = (W2*W2-W0*W0)/W2
      IF (KF.EQ.4) WP = 1.0/WP
      IF (KA.EQ.1) GOTO 30
      PRINT *, 'ENTER PASSBAND RIPPLE OR STOPBAND ATT IN +DB '
      READ *, R1
      PRINT ' '
      IF (KA.EQ.3) WS = WP
      GOTO 30
20    IF (KF.GE.3) GOTO 25
      PRINT *, 'ENTER PASS AND STOPBAND EDGES IN UN-NORMALIZED HZ '
      READ *, FP, FS
      PRINT ' '
      WP = PREWRP(TP*FP)
      WS = PREWRP(TP*FS)
      IF (KF.EQ.2) WP = 1.0/WP
      IF (KF.EQ.2) WS = 1.0/WS
      PRINT *, 'ENTER PASSBAND RIPPLE AND STOPBND. ATT. IN + DB '
      READ *, R1, R2
      PRINT ' '
      GOTO 35
25    PRINT *, 'ENTER F1,F2,F3,F4 FOR BP OR BR FREQS '
      READ *, F1, F2, F3, F4
      PRINT ' '
      W1 = PREWRP(TP*F1)
      W2 = PREWRP(TP*F2)
      W3 = PREWRP(TP*F3)
      W4 = PREWRP(TP*F4)
      W0 = SQRT(W3*W2)
      WP = (W3*W3 - W0*W0)/W3
      WS = (W4*W4 - W0*W0)/W4
      WST = (W0*W0 - W1*W1)/W1
      IF (WST.LT.WS) WS = WST
      IF (KF.EQ.3) GOTO 26
      W0 = SQRT(W1*W4)
      WP = W1/(W0*W0 - W1*W1)
      WS = W2/(W0*W0 - W2*W2)
      WST = W3/(W3*W3 - W0*W0)
      IF (WST.LT.WS) WS = WST
26    PRINT *, 'ENTER PASSBAND RIPPLE AND STOPBAND ATT. IN +DB '
      READ *, R1, R2
      PRINT ' '
      GOTO 35

C-----BUTTERWORTH, CHEBYSHEV, AND ELLIPTIC FILTERS -----
30    CALL ROOTS1
      GOTO 37
35    CALL ROOTS2

C-----HIGHPASS, BANDPASS, AND BAND REJECT XFORMS -----

```

```

37  IF (KF.EQ.1) GOTO 65
    CALL FREQXFM (W0,PR,PI)
    CALL FREQXFM (W0,ZR,ZI)
    IF (KF.EQ.2) GOTO 65
        N2 = N
        N = 2*N
        KOD = 0
65  IF (KAD.EQ.1) GOTO 80

C  ----- DIGITAL BILINEAR TRANSFORM -----
    CALL BLT (N2,SR,PR,PI)
    CALL BLT (N2,SR,ZR,ZI)
    PRINT ' '
    PRINT ',Z PLANE'
    PRINT ' '
    CALL PRNT (N2,PR,PI,ZR,ZI)

C  ----- CASCADE STRUCTURE AND FREQ. RESPONSE -----
80  CALL CASCAD (PR,PI,ZR,ZI,B1,B2,A1,A2,FM)
85  IF (KAD.EQ.2) CALL DFR(KK,B1,B2,A1,A2,FM,STF,SPF)
    IF (KAD.EQ.1) CALL AFR
    CALL VIEW(KK,FM,STF,SPF)
    PRINT ' '
    PRINT ',ENTER 1 FOR EXPANDED FREQ. CALC, 2 FOR A NEW
X FILTER DESIGN, 3 TO ADJUST PASSBAND FREQ., 4 TO END '
    READ ',QUE
        IF (QUE.EQ.1) GOTO 85
        IF (QUE.EQ.2) GOTO 10
        IF (QUE.EQ.3) GOTO 13
    PRINT ',THATS ALL FOLKS'

    END

C  -----END OF MAIN PROGRAM -----

C
C
C
C

C-----BW, CHEBY I&II POLE & ZERO LOCATIONS -----
C
    SUBROUTINE ROOTS1
    DIMENSION PR(20), PI(20), ZR(20), ZI(20)
    COMMON /PARM/R1,R2,WP,WS,N2,N,SR,KA,KAD,KOD,KF
    COMMON /ROOT/PR,PI,ZR,ZI
    ARCSNH(X) = ALOG(X+SQRT(X*X+1))

C
    E = SQRT(10.0**(0.1*R1) - 1)
    IF (KA.EQ.3) E = 1/E
    L = 0
    N2 = (N+1)/2
    KOD = 1
    IF (MOD(N,2).EQ.0) KOD = 0
    IF (KOD.EQ.0) L = 1

```

```

        SM = 1.0
        CM = 1.0
        IF (KAEQ.1) GOTO 10
        V0 = ARCSNH(1/E)/N
        SM = SINH(V0)
        CM = COSH(V0)
10      DO 15 J = 1, N2
            ARG = 1.570796326794897*L/N
            TR = -SM*COS(ARG)
            TI = CM*SIN(ARG)
            ZR(J) = 0.0
            ZI(J) = 1E25
            IF (KAEQ.3) GOTO 11
            PR(J) = WP*TR
            PI(J) = WP*TI
            GOTO 12
11      IF (L.NE.0) ZI(J) = WS/SIN(ARG)
            PR(J) = WS*TR/ (TR*TR + TI*TI)
            PI(J) = WS*TI/ (TR*TR + TI*TI)
12      L = L + 2
15      CONTINUE
        RETURN
        END

C
C
C-----ELLIPTIC FILTER POLE & ZERO LOCATIONS -----
C
        SUBROUTINE ROOTS2
        REAL K, KC, KK, KKC, K1, K1C, KK1, KK1C
        DIMENSION PR(20), PI(20), ZR(20), ZI(20)
        COMMON /ROOT/PR, PI, ZR, ZI
        COMMON /PARM/R1, R2, WP, WS, N2, N, SR, KA, KAD, KOD, KF
        COMMON /ELP1/ SN, CN, DN

C
        E = SQRT(10.0**(0.1*R1)-1)

C
C
C THIS IS NOT FINISHED BUT WHO CARES ABT ELLIPTICAL?
C
15      CONTINUE
        RETURN
        END

C
C
C
C----- PREWARP OF FREQ. BEFORE BLT -----
        FUNCTION PREWRP(WW)
        COMMON /PARM/R1,R2,WP,WS,N2,N,SR,KA,KAD,KOD,KF
        IF (KAD.EQ.1) PREWRP = WW
        IF (KAD.NE.1) PREWRP = 2.0*SR*TAN(WW/2.0/SR)
        RETURN
        END

C
C

```

```

C
C
C----- DIGITAL BILINEAR TRANSFORM BLT -----
C
      SUBROUTINE BLT(N2,SR,R,I)
      REAL SR, R(1), I(1)
C
      A = 2.0*SR
      DO 15 J=1, N2 +1
          TR = R(J)
          TI = I(J)
          IF (ABS(TI).GT.1E15) GOTO 10
          IF (ABS(TR).GT.1E15) GOTO 10
          TD = (A - TR)**2 + TI*TI
          R(J) = (A*A - TR*TR - TI*TI)/ TD
          I(J) = 2.0*A*TI/TD
          GOTO 15
10      R(J) = -1.0
          I(J) = 0.0
15      CONTINUE
          RETURN
          END
C
C
C
C
C----- FREQUENCY TRANSFORMATION -----
C
      SUBROUTINE FREQXFM(W0,PR,PI)
      REAL PR(1), PI(1)
      COMPLEX PC, SC
      COMMON /PARM/R1,R2,WP,WS,N2,N,SR,KA,KAD,KOD,KF
C
      NT = 2*N2+1
      IF (KF.GE.3) GOTO 12
5      DO 10 J=1, N2
          IF (PI(J).GT.1E15) GOTO 7
          PC = CMPLX (PR(J), PI(J))
          SC = 1.0/PC
          PR(J) = -ABS(REAL(SC))
          PI(J) = ABS(AIMAG(SC))
          GOTO 10
7          PR(J) = 0.0
          PI(J) = 0.0
10      CONTINUE
          RETURN
12      DO 14 J=1, N2
          IF (PI(J).GT.1E15) GOTO 13
          PC = CMPLX (PR(J), PI(J))
          IF (KF.EQ.4) PC = 1.0/PC
          SC = (PC - CSQRT(PC*PC - 4*W0*W0))/2.0
          PR(J) = -ABS(REAL(SC))
          PI(J) = ABS(AIMAG(SC))
          SC = (PC + CSQRT(PC*PC - 4*W0*W0))/2.0

```

```

      PR(NT-J) = -ABS(REAL(SC))
      PI(NT-J) = ABS(AIMAG(SC))
      GOTO 14
13      PR(J) = 0.0
      PR(NT-J) = 0.0
      PI(J) = 1E17
      PI(NT-J) = 0.0
      IF (KF.EQ.4) PI(J) = W0
      IF (KF.EQ.4) PI(NT-J) = W0
14      CONTINUE
      RETURN
      END

C
C
C
C
C-----COMPLETE ELLIPTICAL INTERGRAL-----
C
      FUNCTION CE(KC)
C
C STUFF AND THINGS MISSING
C
      RETURN
      END

C
C
C----- ELLIPTICAL FUNCTIONS -----
C
      SUBROUTINE ELP(X,KC)
C
C STUFF MISSING
C
      RETURN
      END

C
C
C----- ARC ELLIPTIC TANGENT -----
C
      FUNCTION ARCSC(U,KC)
C
C STUFF
C
      RETURN
      END

C
C
C----- MODULUS FROM RATIO OF K/K' -----
C
      FUNCTION FK(U)
C
C STUFF
C
      RETURN
      END

```

```

C
C
C
C
C----- PRINT -----
C
      SUBROUTINE PRNT (N2, PR, PI, ZR, ZI)
      DIMENSION PR(20), PI(20), ZR(20), ZI(20)
C
      PRINT *, '#, ZEROS (REAL, IMAG), POLES (REAL,
XIMAG)'
      PRINT *, '
      DO 1 I=1, N2+1
          WRITE (*,10) I,ZR(I), ZI(I), PR(I), PI(I)
1      CONTINUE
10     FORMAT      (I3,4F14.6)
      RETURN
      END
C
C
C
C
C----- CASCADE STRUCTURE PARAMETERS -----
C
      SUBROUTINE CASCAD(PR,PI,ZR,ZI,B1,B2,A1,A2)
      DIMENSION PR(1), PI(1), ZR(1), ZI(1)
      DIMENSION B1(20), B2(20), A1(20), A2(20)
      COMMON /PARM/R1, R2, WP, WS, N2, N, SR, KA, KAD, KOD, KF
      CHARACTER*12 FNCOEF
C
      PRINT *, '
      PRINT *, 'ENTER FILENAME TO CONTAIN As and Bs'
      READ 100,FNCOEF
      OPEN (1, FILE=FNCOEF)
      REWIND (1)
      PRINT *, '
PRINT *, N2, ' CASCADE STAGES, EACH OF THE FORM:'
      PRINT *, 'F(z) = (z*z + B1 z + B2) / (z*z + A1 z + A2)'
      K = 0
      IF ((MOD(N2,2).NE.0).AND.(KF.EQ.3)) K = 1
      J0 = 1
      IF (KOD.EQ.0) GOTO 10
          B1(1) = 1.0
          IF (KF.EQ.2) B1(1) = -1.0
          B2(1) = 0.0
          A1(1) = -PR(1)
          A2(1) = 0.0
          PRINT *, '
          WRITE (*,30) J0, B1(J0), B2(J0), A1(J0), A2(J0)
          WRITE (1,201) J0, B1(J0), B2(J0), A1(J0), A2(J0)
          J0 = 2
10     DO 15 J = J0, N2
          B1(J) = -2.0*ZR(J)
          B2(J) = ZR(J)*ZR(J) + ZI(J)*ZI(J)

```

```

        IF ((J.EQ.1).AND.(K.EQ.1)) B1(J) = 0.0
        IF ((J.EQ.1).AND.(K.EQ.1)) B2(J) = -1.0
        A1(J) = -2.0*PR(J)
        A2(J) = PR(J)*PR(J) + PI(J)*PI(J)
        IF (PI(J).EQ.0) A1(1) = -PR(1) - PR(N2+1)
        IF (PI(J).EQ.0) A2(1) = PR(1)*PR(N2+1)
        WRITE (*,30) J, B1(J), B2(J), A1(J), A2(J)
        WRITE (1,201) J, B1(J), B2(J), A1(J), A2(J)
15      CONTINUE
        RETURN
30      FORMAT ('      ',I3,2F17.10,'      ',2F17.10)
100     FORMAT (A12)
201     FORMAT (I10,2X,E17.10,2X,E17.10,2X,E17.10,2X,E17.10)
        END

C
C
C----- ANALOG FILTER FREQ RESPONSE -----
C
        SUBROUTINE AFR
        PRINT *, 'ANALOG PART NOT FINISHED'
        RETURN
        END

C
C
C----- DIGITAL FILTER FREQ RESPONSE -----
C
        SUBROUTINE DFR(KK,B1,B2,A1,A2,FM,STF,SPF)
        DIMENSION B1(1), B2(1), A1(1), A2(1)
        DIMENSION FM(1)
        COMMON /PARM/R1,R2,WP,WS,N2,N,SR,KA,KAD,KOD,KF
        STF=0
        SPF=0
        STW=0
        SPW=0
        Q=0

C
        PRINT *, ' '
        PRINT *, 'ENTER 1 FOR LOG MAGNITUDE, ELSE 0 '
        READ *, LG
        PRINT *, ' '
        PRINT *, 'ENTER FREQUENCY TO START FREQ. RESPONSE IN Hz '
        READ *, STF
        PRINT *, ' '
        PRINT *, 'ENTER FREQ. TO STOP FREQ. RESPONSE CALC. '
        READ *, SPF
        PRINT *, ' '
        STW=STF/(SR/2)*3.141592654
        SPW=SPF/(SR/2)*3.141592654
        Q = (SPW-STW)/KK
        DO 20 J = 1, KK+1
            W = STW + (Q* (J-1))
            W2 = 2.0 * W
            BR = 1.0
            BI = 0.0

```

```

      AR = 1.0
      AI = 0.0
      IO = 1
      IF (KOD.EQ.0) GOTO 10
      BR = COS(W) + B1(1)
      BI = SIN(W)
      AR = COS(W) + A1(1)
      AI = SIN(W)
      IO = 2
10      DO 15 I = IO, N2
      BRT = COS(W2) + B1(I)*COS(W) + B2(I)
      BIT = SIN(W2) + B1(I)*SIN(W)
      ART = COS(W2) + A1(I)*COS(W) + A2(I)
      AIT = SIN(W2) + A1(I)*SIN(W)
      BRS = BR*BRT - BI*BIT
      BI = BR*BIT + BI*BRT
      BR = BRS
      ARS = AR*ART - AI*AIT
      AI = AR*AIT + AI*ART
      AR = ARS
15      CONTINUE
      FM(J) = SQRT((BR*BR + BI*BI)/(AR*AR + AI*AI))
      IF (LG.EQ.0) GOTO 20
      IF (FM(J).EQ.0) FM(J) = 1E-5
      FM(J) = 20*LOG10(FM(J))
20      CONTINUE
      RETURN
      END
C
C----- OUTPUT FREQUENCY RESPONSE -----
C
      SUBROUTINE VIEW(KK,FM,STF,SPF)
      DIMENSION FM(1)
      CHARACTER*12 FNAME
      REAL INC
C
      PRINT *, 'ENTER FILENAME TO STORE FREQUENCY RESPONSE '
      READ 101, FNAME
      PRINT ' '
      OPEN (2,FILE=FNAME)
      REWIND(2)
      RANG= SPF-STF
      INC = RANG/KK
      DO 10 J = 1, KK+1
      F = STF + INC*(J-1)
      WRITE (2,100) F,FM(J)
10      CONTINUE
100     FORMAT ( F8.3,',',10X,E15.8)
101     FORMAT (A12)
      RETURN
      END

```


Appendix IV-B

Filter Response Curves

The frequency response of the 12 channel bandpass filter bank is shown in Figure 62. Note the 12-bit AIB board was used as the analog interface to the TMS320c25, hence the approximate 72 dB noise floor, shown in the figure.

Figure 62 Bandpass Filter Bank Frequency Response

Appendix IV-C **TMS320c25 Peak Picking Code Sample**

The code listing below illustrates the peak picking program for the highest frequency bandpass filter (CH14). The code for the other channels is similar and follows in-line after this section. The results of the peak picker are stored in memory location HI#, (where # is the filter number) and are used to calculate the principal component amplitudes.

```

.....
*
*
* START OF PEAK PICKING CODE
*
* AR5,AR6,AR7 ARE USED HERE GLOBALLY i.e. CAN'T ALTERED ELSEWHERE
*
* AR5 = 56, AR6 = 20, AR7 = 7 AND ARE NUMBER OF SAMPLES IN PP WINDOW
*
.....
***pp
*
      LDPK 6
      LAC  VLP14,15      * LOAD BPF14 OUTPUT
      ABS                      * ABSOL VALUE
      SUB  HI14,15      * SUBTRACT HIGHEST PREV. VALUE
      BLZ  NC14          * IF LESS THAN 0-> CHECK IF DONE
*
      ADD  HI14,15      * ELSE STORE CURRENT VALUE AS HIGHEST
      SACH HI14,1
*
NC14  LARP 7            * LOAD SAMPLE COUNTER AR7
      BANZ N13,*        * IF NOT DONE GO TO NEXT BPF OUTPUT
*
      DMOV HI14          * ELSE SAVE HIGHEST VALUE OF PREVIOUS
      ZAC                      * GROUP OF SAMPLES
      SACL HI14
*
      LAC  HIZ14,12
      SACH HIZ14
*
*
N13  LAC  VLP13,15      * AS ABOVE FOR THE OTHER 12 CHANNELS

```

Appendix IVD Logarithm Calculation Details

The code listing for the LOGAP2.FOR Fortran program and the TMS320c25 linear to log conversion program are listed below. The LOGAP2 program operates on the following principle:

For a portion of the curve $20\log(x)$ as shown, a linear line segment may be selected to approximate it to a given error, e_1 , as shown in Figure 63.

Figure 63 Linear Approximation of a Log Curve Segment

The starting point on the linear approximation segment (for the log section starting at $x=1$) is $(x_1, y_1) = (1, e_1)$, where e_1 is the allowable error. Over the length of the log curve approximated by the linear segment, there are two other locations for maximum error, e_2 and e_3 . The equation of the line segment using point (x_1, y_1) is:

$$y_1 = mx_1 - b, \quad \text{but } x_1 = 1$$

$$\therefore y_1 = m - b$$

$$\text{and } b = m - y_1 \quad (1)$$

The maximum midpoint error, e_2 occurs at x_2 , when the slope of the log curve, y and the linear segment, y_1 are equal.

Slope of log curve:

$$\begin{aligned} m &= \partial/\partial x(20\log x_2) \\ m &= 20(\log e)\partial/\partial x(x_2) \\ m &= 8.685889/x_2 \end{aligned} \quad (2)$$

Substituting into linear segment equation:

$$\begin{aligned} y_2 &= mx_2 - b \\ y_2 &= (8.685889/x_2)x_2 - b \\ y_2 &= 8.685889 - b \end{aligned} \quad (3)$$

Value of the error at x_2 : $e_2 = 20\log x_2 - y_2 \leq e_1$

Substituting (3) in :

$$\begin{aligned} &= 20\log x_2 - 8.685889 + b \leq e_1 \\ \therefore 20\log x_2 &\leq e_1 - b + 8.685889 \\ 20\log x_2 + b &\leq e_1 + 8.685889 \end{aligned} \quad (4)$$

Substituting (1) into (4):

$$20\log x_2 + 8.685889/x_2 - y_1 - 8.685889 \leq e_1 \quad (5)$$

Knowing the starting point of the log curve, ($y=0$ in Figure 63), and the allowable error, e_1 , $y_1 = 0 + e_1$, and equation (5) can be solved by iteration. The LOGAP2.FOR program does this by taking a user input guess ($G2$) for the $X2$ value and the user input iteration step size (S), and solves equation (5) for the value of $G2$ which brings the error as close as possible to the limit.

$G2$ is then assigned to x_2 , and equations (1) and (2) may be solved for b and m , respectively. Having the slope and intercept values for the linear approximation segment, the end point of the line segment, x_3 may be determined, again by iteration, of equation (6):

$$mx_3 + b - 20\log x_3 \leq e_1 \quad (6)$$

The LOGAP2 program allows a different step size ($S1$) to be used in this iteration for increased resolution. LOGAP2.FOR then saves the line segment number, slope (m), intercept (b) and end point (x_3 , y_3) values to a file and allows the user to continue with the next line segment or to exit the program.

LOGAP2.FOR:

```

C
C LOGAP2          88 12 08
C
C This program calculates a piecewise linear approximation
C to the log base 10 function, starting at a linear value of 1.
C Input parameters are: starting point X1 and Y1
C                      allowed error - E
C                      first guess for finding point of max error
C                      btwn end points - G2
C                      step size of iterations - S
C
C Output: to file of slopes and y-intercepts of the sections
C         - start and end values of the line segments
C
C
C
C DIMENSION MF(30),XF(30),YF(30),BF(30)
C REAL X1,Y1,TM,G2,E1,S,MX,TB,START,MF,XF,YF,BF
C REAL X2,M,B,END,G3,X3,S1
C REAL Y3S,Y3LG,DIFF,Y2S,Y2LG,DIFF2
C INTEGER J
C CHARACTER*12 FNAME
C
C PRINT *, 'ENTER THE STARTING X-VALUE: '
C READ *, X1
C PRINT *, ' '
C PRINT *, 'ENTER THE STARTING Y-VALUE W/O ERROR: '
C READ *, Y1
C PRINT *, ' '
C PRINT *, 'ENTER THE ALLOWED ERROR IN dB: '
C READ *, E1
C PRINT *, ' '
C
C Y1 = Y1 + E1
C
C PRINT *, 'ENTER THE FILE NAME TO STORE START,M,B,END: '
C READ 10, FNAME
C 10  FORMAT (A12)
C PRINT *, ' '
C OPEN (1, FILE=FNAME)
C REWIND (1)
C
C J = 1
C WRITE (1,11) X1,Y1
C 11  FORMAT (' 0 , 0 , ', E17.10, ',', 2X, E17.10)
C
C 14 PRINT *, 'ENTER THE STEP SIZE FOR MID-PT. ITERATIONS: '
C READ *, S
C PRINT *, ' '
C
C PRINT *, 'ENTER THE STEP SIZE FOR END-PT. ITERATIONS: '

```

```

      READ *,S1
      PRINT *, ' '
C
15  TM = 20 * LOG10(2.7182818)
C
      MX = TM * X1
C
      IF (X1.GT.850) S = 0.1
C
C
16  PRINT      *, 'ENTER THE FIRST GUESS FOR MAX MID ERROR X: '
      READ *,G2
      PRINT *, ' '
C
      START = 20*LOG10(G2) - TM - Y1 + TM/G2
C
      WRITE      (*,17) START
17  FORMAT      (2X,F10.4)
      PRINT *, 'ENTER 1 TO CONTINUE OR 2 FOR NEW GUESS: '
      READ *,NY
      IF (NY.EQ.1) GOTO 50
      IF (NY.EQ.2) GOTO 16
C
50  IF (START.LT.E1) GOTO 100
C
90  G2 = G2 - S
      START = 20*LOG10(G2) - TM - Y1 + MX/G2
      WRITE (*,91) G2,START
91  FORMAT      (10X,F14.5,10X,F12.6)
      IF (START.GT.E1) GOTO 90
      X2 = G2
      GOTO 200
C
100 G2 = G2 + S
      START = 20*LOG10(G2) - TM - Y1 + MX/G2
      WRITE (*,91) G2,START
      IF (START.LT.E1) GOTO 100
      X2 = G2 - S
C
200 M = TM / X2
      B = Y1 - M*X1
C
C
      WRITE      (*,205)M,B,X2,Y1
205 FORMAT      (F6.2,2X,F6.2,2X,F10.4,2X,F10.4)
      PRINT *, ' '
C
      PRINT *, 'THE X2 VALUE IS: '
      WRITE      (*,210) X2
210 FORMAT      (' ',F17.10)
220 PRINT *, ' '
      PRINT *, 'ENTER GUESS FOR X3 VALUE: '
      READ *,G3
C

```

```

280  END = M*G3 + B - 20*LOG10(G3)
      WRITE      (*,285)G3,END
285  FORMAT      (2X,F12.3,10X,F15.6)
      PRINT *,'ENTER 1 TO CONTINUE, 2 FOR NEW GUESS:'
290  READ  *,YN
      IF (YN.EQ.1) GOTO 295
      IF (YN.EQ.2) GOTO 220
      GOTO 290

C
295  IF (END.LT.E1)      GOTO 400
C
300  G3 = G3 - S1
      END = M*G3 + B - 20*LOG10(G3)
      WRITE (*,91) G3,END
      IF (END.GT.E1)      GOTO 300
      X3 = G3
      GOTO 500

C
400  IF (END.LT.0) GOTO 450
410  G3 = G3 + S1
      END = M*G3 + B - 20*LOG10(G3)
      WRITE (*,91) G3,END
      IF (END.LT.E1) GOTO 410
      X3 = G3 - S1
      GOTO 500

C
450  PRINT *,'ENTER HIGHER GUESS FOR X3 VALUE:'
      READ  *,G3
      GOTO 280

C
C
C
500  Y3S = M*X3 + B
      Y3LG = 20*LOG10(X3)
      DIFF = Y3S - Y3LG

C
      Y2S = M*X2 + B
      Y2LG = 20*LOG10(X2)
      DIFF2 = Y2LG - Y2S

C
      PRINT *, '
      PRINT      *, 'THE MIDPOINT RESULTS Y2LG, Y2S, DIFF. ARE:'
      PRINT *, '
      WRITE      (*,510) Y2LG,Y2S,DIFF2
510  FORMAT      (10X,F17.10,5X,F17.10,5X,F17.10)
      PRINT *, '

C
      PRINT *, 'THE END POINT RESULTS Y3LG, Y3S AND DIFF ARE:'
      PRINT *, '
      WRITE      (*,520) Y3LG,Y3S, DIFF
520  FORMAT      (10X,F17.10,5X,F17.10,5X,F17.10)
      PRINT *, '

C

```



```

C      MF(J) = M
      BF(J) = B
      XF(J) = X3
      YF(J) = Y3LG + E1
C
      PRINT *,THE EQUATION AND END POINT OF THE LINE IS:
      PRINT *,
      WRITE      (*,530)  J,MF(J),BF(J),XF(J),YF(J)
      WRITE      (1,540)  MF(J),BF(J),XF(J),YF(J)
530    FORMAT      (2X,'J= ',I2,2X,'Y = ',F8.5,'*X +',F6.2,' X3 =',E11.4,
      X2X,'Y3 = ',F6.2)
540    FORMAT      (E17.10,',',E17.10,',',E17.10,',',E17.10)
C
C
      PRINT *,DO YOU WISH TO CONTINUE YES=1, NO=0
      READ  *,YN
      IF (YN.EQ.0) GOTO 1000
C
      X1 = X3
      Y1 = YF(J)
C
      J = J+1
C
      WRITE      (*,600)  S,S1
600    FORMAT      (2X,'CURRENT STEP SIZES ARE: S=',F5.3,' S1=',F5.3)
C
      PRINT      *,ENTER 1 TO CHANGE STEP SIZES, 0 TO CONT.:
      READ  *,YNS
      IF (YNS.EQ.0) GOTO 15
      GOTO 14
C
1000  J = J + 1
      WRITE (*,1010) J
1010  FORMAT      (' THE NUMBER OF LINEAR SEGMENTS IS:', I3)
      PRINT *,THATS ALL FOLKS
C
      END

```

TMS320c25 Implementation:

```

*.....*
*
* LOG CALC. SECTION FOLLOWS      USES AR2,AR3,AR4 LOCALLY
*
*.....*
***LOG
*
* START OF LOG CALCULATION
*
LOG  LARK  0,2  *SET AR0 TO INC. VALUE OF 2

```

```

    LARK 3,1  * PASS THRU LOG LOOP TWICE UNDER CTRL OF AR3
    LARP 2
    .
    .
    LDPK 0
    SPM 3  * SET P REG SHIFTER TO RIGHT 6BITS
    .
    * AR2 = >30F TO START AND IS THE ADDRESS OF THE PP DATA WHICH IS TO BE
    * CONVERTED TO THE LOG.
    * AR4 = >350 AND IS THE START ADDRESS OF THE STORAGE LOCATIONS TO CONTAIN
    * THE LOG. VALUE OF THE ENERGY
    .
LGLP  LAC  *
      RPTK 2  * determine which linear segment should be used
      SFR
      BLEZ ZERO
      SUBK 179
      BLZ LOW
      SBLK 3042
      BLZ T1227
      SBLK 5238
      BLZ CALJ
      SBLK 13761
      BLZ CALK
      B CALCL
T1227 LAC  *
      SBLK 468
      BLZ CALG
      SBLK 759
      BGEZ CALI
      B CALH
    .
LOW  LAC  *
      SUBK 11
      BGEZ T27
      LAC  *
      SUBK 5
      BLZ CALAB
      B CALC
T27  SUBK 16
      BLZ CALD
      SUBK 42
      BGEZ CALF
    .
    . . .
    .
      LAC  BE  * LOAD B VALUE FOR E AMPL. RANGE
      LT  *  * LOAD PP VALUE INTO T REGISTER
      MPY ME  * MULTIPLY PP VALUE BY SLOPE, M FOR E AMP. RANGE
      APAC  * ADD M*PP TO B VALUE
      MAR *0+,4  * INC. PP ADDRESSING BY 2, LOAD LOG STORAGE ADDR.
      SACL *+,0,3  * STORE LOG, INC ADDR., LOAD LOG LOOP COUNTER AR3
      B DONE
    .
ZERO MAR *0+,4

```

	SACL	*+,0,3	
	B	DONE	
.			
CALAB	LDPK	6	
	MAR	*0+,2	* INC PP. ADDRESSING BY 2, POINT TO AR2
	SAR	2,TEMP	* STORE CURRENT PP ADDRESS IN TEMP
	ADDK	>64	* ADD OFFSET TO ACCUM. IN ORDER TO OBTAIN LOG ADDR.
	SACL	ADDR	* STORE ADDR. OF LOG VALUE IN ADDR
	LAR	2,ADDR	* LOAD AR2 WITH ADDR. OF CORRECT LOG VALUE
	LAC	*-,0,4	* LOAD ACCUM WITH LOG VALUE, POINT TO STORAGE ADDR.
	SACL	*+,0,3	* STORE LOG VALUE, LOAD AR3 LOG LOOP COUNTER
	LAR	2,TEMP	* RELOAD PP ADDR. INTO AR2
	LDPK	0	
	B	DONE	
.			
.			
.			
CALC	LAC	BC	
	LT	*	
	MPY	MC	
	APAC		
	MAR	*0+,4	
	SACL	*+,0,3	
	B	DONE	
.			
CALD	LAC	BD	
	LT	*	
	MPY	MD	
	APAC		
	MAR	*0+,4	
	SACL	*+,0,3	
	B	DONE	
.			
CALF	LAC	BF	
	LT	*	
	MPY	MF	
	APAC		
	MAR	*0+,4	
	SACL	*+,0,3	
	B	DONE	
.			
CALG	LAC	BG	
	LT	*	
	MPY	MG	
	APAC		
	MAR	*0+,4	
	SACL	*+,0,3	
	B	DONE	
.			
CALH	LAC	BH	
	LT	*	
	MPY	MH	
	APAC		
	MAR	*0+,4	

```

      SACL  *+,0,3
      B     DONE
.
CALI  LAC   BI
      LT    *
      MPY   MI
      APAC
      MAR   *0+,4
      SACL  *+,0,3
      B     DONE
.
CALJ  LAC   BJ
      LT    *
      MPY   MJ
      APAC
      MAR   *0+,4
      SACL  *+,0,3
      B     DONE
.
CALK  LAC   BK
      LT    *
      MPY   MK
      APAC
      MAR   *0+,4
      SACL  *+,0,3
      B     DONE
.
CALCL LAC   BL
      LT    *
      MPY   ML
      APAC
      MAR   *0+,4
      SACL  *+,0,3
.
DONE  BANZ  LGLP,*-,2
.

```

Appendix V Tactile Psychophysical Testing

Refer to Figure 41, in Chapter 4 for the experimental set-up.
The three programs and test implemented are:

- i) amplitude threshold - TACMOA.BAS
- ii) maximum comfortable level - TACMAX.BAS
- and iii) selection of an active transducer from an array - TACSEL.BAS.

Due to time constraints, only the first two tests were performed. The software was written in the BASIC language on a Digital Equipment Corp. (DEC) MINC-11 computer, running the RT-11 operating system. The listings of the three programs and flow diagrams follow.

TACMOA.BAS

```
10 REM set up array to hold threshold values:
20 DIM T(10)
21 PRINT 'Input filename to store data'; \ INPUT F$
23 PRINT ''
30 REM Clear IEEE488 bus:
40 IEEE_BUS_CLEAR
50 REM Set time out of the bus:
60 INSTR_TIME_LIMIT(0)
70 ALL_INSTR_CLEAR
78 REM Initialize function generator via IEEE488 command
79 REM Set to 0 vpp, 250 Hz:
80 SEND("A0F250Q0B0N0C0P0",2)
81 C=1
82 Z=0
83 REM Remind me to set the pot. dial to the centre before
84 REM start of expt.-> prog. tells me when it's centred:
85 PRINT 'SET DIAL TO CENTRE'
86 REM Get pot voltage to see if it is at centre:
87 AIN(V1) \ V1=V1+5.12
88 IF V1>=5 THEN PRINT 'OK' \ GO TO 100
89 PRINT 'KEEP GOING'
90 GO TO 87
100 RANDOMIZE
110 REM Get random number for first pot setting:
120 R=RND
130 REM Get pot voltage value:
140 AIN(V1)
150 REM Scale pot value to 0 thru 10.24 volts:
160 V1=V1+5.12
170 REM Set the func. gen. output value to start the test:
```

```

171 A=5.00000E-03*V1+.165*R
172 IF V1<2.5 THEN A=1.25000E-03*V1+.012*R
176 IF V1>7.5 THEN A=.01*V1+.1*R
210 REM Set up message string to func. generator:
220 M$='A'+STR$(A)+';P1;l;'
225 PRINT 'BEGIN TRIAL 'C
229 REM Send IEEE488 message to function generator:
230 SEND(M$,2)
240 REM Get potent. voltage value:
250 AIN(,V2)
255 V2=V2+5.12
259 REM Adjust func. gen. amplitude to follow pot. changes:
260 IF V2>=V1 THEN A=((V2-V1)/18.725)+A \ GO TO 275
270 IF V2<V1 THEN D=(V1-V2)/18.725
271 IF D>A THEN D=A
273 A=A-D
275 V1=V2
276 V2=0
280 REM Send new amplitude value to func. gen:
285 SEND('A'+STR$(A)+';l',2)
290 REM Check if time for adjustment has expired
291 REM Allow 60 times thru the loop:
295 Z=Z+1
300 IF Z<60 THEN 250
305 Z=0
310 PRINT 'Is this your threshold ? (Y/N)'; \ INPUT B$
320 IF B$>'P' THEN 325
321 AIN(V) \ V=V+5.12
322 IF V<.1 THEN 100
323 IF V>10 THEN 100
324 GOTO 220
325 SEND('P010',2)
326 IF A<=1.500000E-03 THEN PRINT 'THRESHOLD IS 0, try again' \ GO TO 220
327 PRINT ''
330 T(C)=A*2
340 C=C+1
345 REM Check if done all 10 trials, if not go thru loop again:
350 IF C<=10 THEN 100
355 ALL_INSTR_CLEAR
360 PRINT 'YOU ARE FINISHED, HERE ARE YOUR THRESHOLD VALUES:'
361 G$='SY1:' + SEG$(F$,1,8)
365 OPEN G$ FOR OUTPUT AS FILE #1
370 FOR I=1 TO 10
380 PRINT 'T(';I;')=',T(I)
381 PRINT #1,'T(';I;')=',T(I)
385 M=M+T(I)
390 NEXT I
391 M=M/10
392 PRINT 'THE AVERAGE THRESHOLD VALUE IS:',M
393 PRINT #1,'THE AVERAGE THRESHOLD VALUE IS:',M
394 CLOSE #1
400 END

```

TACMAX.BAS

```

10  REM set up array to hold maximum values:
20  DIM T(10)
30  PRINT 'Input filename to store data'; \ LINPUT F$
40  PRINT ''
50  REM Clear IEEE488 bus:
60  IEEE_BUS_CLEAR
70  REM Set time out of the bus:
80  INSTR_TIME_LIMIT(0)
90  ALL_INSTR_CLEAR
98  REM Initialize function generator via IEEE488 command
99  REM Set to 0 vpp, 250 Hz:
100 SEND("A0F250Q0B0N0C0P0",2)
110 C=1
120 T=100
130 Z=0
140 M=0
150 X=0
151 REM Remind me to set the pot. dial to the centre before
152 REM start of expt.-> prog. tells me when it's centred:
160 PRINT 'SET DIAL TO CENTRE'
169 REM Get pot voltage to see if it is at centre:
170 AIN(,V1) \ V1=V1+5.12
180 IF V1>=5 THEN PRINT 'OK' \ GO TO 100
190 PRINT 'KEEP TURNING.....' \ GO TO 170
191 PRINT 'STOP'
200 RANDOMIZE
210 REM Get random number for first pot setting:
220 R=RND
229 REM Get pot voltage value:
230 AIN(,V1)
239 REM Scale pot value to 0 thru 10.24 volts:
240 V1=V1+5.12
250 REM Set the func. gen. output value to start the test:
260 A=2*R+X
270 REM Set up message string to function generator:
280 M$='A'+STR$(A)+';P1;I;'
290 PRINT 'BEGIN TRIAL 'C
299 REM Send IEEE488 message to function generator:
300 SEND(M$,2)
310 REM Get potent. voltage value:
320 AIN(,V2)
330 V2=V2+5.12
339 REM Adjust func. gen. amplitude to follow pot. changes:
340 IF V2>=V1 THEN A=((V2-V1)/2)+A \ GO TO 380
350 IF V2<V1 THEN D=(V1-V2)/3
360 IF D>A THEN D=A
370 A=A-D
379 REM Check if Amplitude is greater than allowed by vibrator:
380 IF A>4 THEN 550

```

```

390 V1=V2
400 V2=0
409 REM Send new amplitude value to func. gen:
410 SEND('A'+STR$(A)+';',2)
420 REM Check if time for adjustment has expired
430 REM Allow 60 times thru the loop:
440 Z=Z+1
450 IF Z<60 THEN 320
460 Z=0
470 PRINT 'Is this your max level ? (Y/N)'; \ INPUT B$
480 IF B$>'P' THEN 630
490 AIN(V) \ V=V+5.12
499 REM Check if pot. setting is at either end of range, if so restart trial:
500 IF V<.1 THEN 150
510 IF V<10 THEN 280
520 X=1.5
529 REM If user not yet at max. level, continue:
530 GOTO 160
540 REM Come to 550 if vibrator maximum is reached:
550 IF A>4 THEN PRINT 'MAXIMUM vibrator level reached'
559 REM Set the vibrator to the max. (8vpp) to let subject feel the max value:
560 A=4
570 SEND('A'+STR$(A)+';',2)
580 PRINT 'This is the maximum output level of the vibrator'
590 M=1
600 T=100
610 T=T-1
620 IF T>0 THEN 610
630 PRINT ''
631 SEND('P010',2)
639 REM Convert A value from volts peak to volts peak-peak and store in T
640 T(C)=A^2
650 C=C+1
659 REM Check if done all 10 trials, if not go thru loop again:
660 IF C<=10 THEN 690
670 IF M=0 THEN 200
680 GOTO 160
690 ALL_INSTR_CLEAR
700 PRINT 'YOU ARE FINISHED, HERE ARE YOUR MAXIMUM VALUES:'
710 G$='SY1:' + SEG$(F$,1,8)
720 OPEN G$ FOR OUTPUT AS FILE #1
730 FOR I=1 TO 10
740 PRINT 'T(';I;')=';T(I)
750 PRINT #1,'T(';I;')=';T(I)
760 M=M+T(I)
770 NEXT I
780 M=M/10
790 PRINT 'THE AVERAGE MAXIMUM VALUE IS:';M
800 PRINT #1,'THE AVERAGE MAXIMUM VALUE IS:';M
810 CLOSE #1
820 END

```


TACSEL

```

10  REM set up array to hold POSITION data:
20  DIM P(32)
28  REM D is data word that controls which vibrator is on
29  REM Disable all vibrators:
30  D=4
40  DOUT(D)
50  PRINT 'Input filename to store data'; \ LINPUT F$
60  PRINT ''
70  REM Clear IEEE488 bus:
80  IEEE_BUS_CLEAR
90  REM Set time out of the bus:
100 INSTR_TIME_LIMIT(0)
110 ALL_INSTR_CLEAR
118 REM Initialize function generator via IEEE488 command
119 REM Set to 0 vpp, 250 Hz:
120 SEND("A0F250Q0B0N0C0P0",2)
130 PRINT 'PLEASE ENTER THE AMPLITUDE OF THE STIMULUS'; \ LINPUT A
140 IF A<4 THEN 170
150 PRINT 'AMPLITUDE MUST BE LESS THAN 4 Vp'
160 PRINT 'TRY AGAIN' \ GO TO 130
170 C=1
180 T=0
190 Z=0
200 O=0
210 H=0
230 REM Set up message string to function generator:
240 M$='A'+STR$(A)+';P1;l;'
250 SEND(M$,2)
260 RANDOMIZE
270 REM Get random number to select 'ON' vibrator:
280 R=RND
289 REM Select vibrator #4:
290 IF R>.75 THEN 440
299 REM Select vibrator #3:
300 IF R>.5 THEN 400
309 REM Select vibrator #2:
310 IF R>.25 THEN 360
317 REM Else select vibrator #1
318 REM Test if vibrator #1 has been used more than 4 times
319 REM If so, then go back to rnd # to select a different vib.:
320 IF Z>=4 THEN 260
321 REM Set vibrator #0 to be activated:
330 D=0
340 Z=Z+1
350 GO TO 470
360 IF O>=4 THEN 260
370 D=1
380 O=O+1
390 GO TO 470

```

```

400 D=2
410 IF T>4 THEN 260
420 T=T+1
430 GO TO 470
440 D=3
450 IF H>=4 THEN 260
460 H=H+1
469 REM Save the vibrator # selected by the program in P:
470 P(C)=D
471 REM Enable a vibrator by sending a digital word out the parallel port
472 REM to a DG528 multiplexor:
475 DOUT(D)
480 PRINT 'BEGIN TRIAL 'C
490 PRINT ''
500 PRINT 'ENTER NUMBER CORRESPONDING TO THE VIBRATOR THAT IS ON'
510 INPUT N
511 IF N>4 THEN 500
512 IF N<0 THEN 500
513 IF N=0 THEN END
520 PRINT 'YOU HAVE SELECTED VIBRATOR #'N
530 PRINT 'IS THIS CORRECT? Y/N'; \ INPUT Q$
540 IF Q$ < 'P' THEN 500
550 SEND ('P010',2)
560 D=4
570 DOUT(D)
580 P(C+16)=N
590 PRINT ''
610 C=C+1
620 IF C<16 THEN 260
650 ALL_INSTR_CLEAR
660 PRINT 'YOU ARE FINISHED, HERE ARE YOUR CHOICES:'
670 G$='SY1:' + SEG$(F$,1,8)
680 OPEN G$ FOR OUTPUT AS FILE #1
690 FOR I=1 TO 16
700 PRINT 'P('I:;')='P(I)
750 PRINT #1,'P('I:;')='P(I)
770 NEXT I
780 CLOSE #1
800 END

```

References

- Aguilera, S. et al. (1986). "Speech-Analysis-Based Devices for Diagnosis and Education of Speech and Hearing Impaired People", I.C.A.S.S.P., 641-648.
- Ahmed, N. and Rao, K.R. (1975). Orthogonal Transforms for Digital Signal Processing, Springer-Verlag, New York.
- Ainsworth, W.A. (1976). Mechanisms of Speech Recognition, Pergamon Press, Toronto, pp 12-15.
- Allerhand, Michael (1987). Knowledge-Based Speech Pattern Recognition, Kogan Page, London, U.K.
- Alles, David (1970). "Information Transmission by Phantom Sensations", IEEE Trans. on Man-Machine Systems, Vol. MMS-11, No. 1, 85-91.
- American Soc. for the Hard of Hearing (1943). "New Aids and Materials for Teaching Lipreading.", ASHH, Washington D.C., John Walters publishers, Amsterdam.
- Analog Interface Board User's Guide (1987). Texas Instruments Inc.
- Anderson, A.B. and Munson, W.A. (1951). "Electrical excitation of nerves in the skin at audio frequencies", JASA, 23, 135-139.
- Anderson, T.W. (1963). "Asymptotic Theory for Principal Component Analysis", 122-148.
- Anderson, T.W. (1971). An Introduction to Multivariate Statistical Analysis, Wiley Publ., 451-477.
- Angelocci, A.A., (1962). "Some observations on speech of the deaf", Volta Review, 64, 403-405.
- Angelocci, A.A., Kopp, G.A., and Holbrook, A., (1964). "The Vowel Formants of Deaf and Normal-Hearing Eleven- to Fourteen-Year-Old Boys JSHD, Vol. 29, No. 2, 156-170.
- Audiological Engineering Corp., (1986). Promotional literature.
- Baum, Shari R. and Blumstein, Sheila E. (1987). "Preliminary observations on the use of duration as a cue to syllable-initial fricative consonant voicing in English", JASA 82 (3), 1073-1076.

Behrens, Susan and Blumstein, Sheila E. (1988). "On the role of the amplitude of the fricative noise in the perception of the place of articulation in voiceless fricative consonants", *JASA*, 84 (3), 861-867.

Berger, K.W. (1972). "Three experiments in speechreading", *J. Commun. Disorders* 5, 280-285.

Bernstein, L.E. (1986). "Speech Training Devices For Profoundly Deaf Children", *I.C.A.S.S.P.*, 633-636. MoliEnergy (1989). Promotional literature and seminar, Ottawa, March.

Bernstein, Lynne E. (1988). "Tactile Aids to Lipreading: Current Status and Future Prospects", *ICAART*, Montreal, 218-219.

Bernstein, Lynne E., Eberhardt, Silvio P., and Demorest, Marilyn E. (1989). "Single-channel vibrotactile supplements to visual perception of intonation and stress", *JASA*, 85, (1), 397-405.

Bernstein, Lynne E., Schechter, Miriam B., and Goldstein, Moise H. (1986). "Child and adult vibrotactile thresholds for sinusoidal and pulsatile stimuli", *JASA*, 80(1), 118-123.

Binnie, Carl A., Jackson, Pamela L., and Montgomery, Allen A. (1976). "Visual Intelligibility of Consonants: A Lipreading Screening Test With Implications For Aural Rehabilitation.", *JSHD*, XLI, 530-539.

Blamey, P.J. and Clark, G.M. (1985). "A wearable multiple-electrode electrotactile speech processor for the profoundly deaf", *JASA*, 77(4), 1619-1620.

Blamey, P.J. and Clark, G.M. (1987). "Psychophysical studies relevant to the design of a digital electrotactile speech processor", *JASA*, 82(1), 116-125.

Blamey, P.J., Cowan, R.S.C., Alcantara, J.I. and Clark, G.M. (1988). "Phonemic Information Transmitted by a Multichannel Electrotactile Speech Processor", *JSHR*, Vol. 31, 620-629.

Bliss, J.C. and Linvill, J.G. (1967). "A Direct Translation Reading Aid" in R. Dulton (ed.), Proceedings of the International Conference on Sensory Devices for the Blind, St. Dunstan's, London.

Blumstein, Sheila E. and Stevens, K.N. (1979). "Acoustic invariance in speech production: Evidence from measurements of the spectral characteristics of stop consonants", *JASA*, 66, 1001-1017.

Boothroyd, Arthur (1984). "Experiments with a Wearable Tactile Pitch Display", Proceedings of the 2nd International Conference on Rehabilitation Engineering.

Boothroyd, Arthur (1985). "A Wearable Tactile Intonation Display for the Deaf", IEEE Trans. on ASSP, Vol. ASSP-33, No. 1, 111-117.

Boothroyd, Arthur and Hnath-Chisolm, Theresa (1987). "Portapitch II: A Multichannel Tactile Display of Voice Fundamental Frequency", Resna 10th Annual Conference, San Jose, Calif., 425-427.

Boothroyd, Arthur, Hnath-Chilsom, Theresa, and Kishon Rabin, Liat (1988). "Single and Multichannel Tactile Displays of F0 as Aids to Lipreading: Speech Pattern Contrast Perception", ICAART, Montreal, 226-227.

Boothroyd, Arthur, Hnath-Chisolm, Theresa, and Hanin, Laurie (1986a). "Tactile Detection of Pitch Falls: Single Channel Stimulation at the Sternum", Preliminary Report #TSL14, CUNY.

Ibid, (1986b). "Tactile Detection of Pitch Falls: Part 2", Preliminary Report #TSL11, CUNY.

Ibid, (1985). "Tactile Detection of Pitch Falls", Preliminary Report #TSL6, CUNY.

Boyd et al. (1986). "A Spectral Visual Feedback System for the Hearing Impaired", IEE Conference on Speech Input/Output Techniques and Applications, IEE Conf. Publ. #258.

Breeuwer, M and Plomp, R. (1986). "Speechreading supplemented with auditorily presented speech parameters", J. Acoustical Soc. Am. 79 (2), 686-691.

Breeuwer, M. and Plomp, R. (1986). "Speechreading supplemented with auditorily presented speech parameters", JASA, 79 (2), 686-691.

Breeuwer, M., and Plomp, R. (1984). "Speechreading supplemented with frequency-selective sound-pressure information", JASA, 76, (3), 686-691.

Brooks, P.L. and Frost, B.J. (1983). "Evaluation of a tactile vocoder for word recognition", JASA, 74, (1), 34-39.

Brooks, P.L., Frost, B.J., Mason, J.L. and Gibson, D.M. (1987). "Word and Feature Identification by Profoundly Deaf Teenagers Using the Queen's University Tactile Vocoder", JSHR, Vol. 30, 137-141.

- Brooks, P.L., Frost, B.J., Mason, J.L., and Chung, K. (1985). "Acquisition of a 250-word vocabulary through a tactile vocoder", *JASA*, 77, (4), 1576-1579.
- Brooks, Patricia L. (1984). "Comprehension of Speech by Profoundly Deaf and Normal-Hearing Subjects Using the Queen's University Tactile Vocoder", Unpublished Ph.D. thesis, Queen's University, Kingston, Canada.
- Brown, R.L., Sporn, R.A., Schmitt, K., and Solomon, A. (1966). "Stimulus parameter considerations and individual differences in cutaneous sensitivity to electropulse stimulation", *Percept. Mot. Skills*, 23, 1215-1222.
- Bustamante, Diane K. (1986). "Principal Component Amplitude Compression of Speech for the Hearing Impaired", Unpublished Doctoral Thesis, Massachusetts Institute of Technology.
- Calvert, D.R., (1962). "Speech sound duration and the surd-sonant error.", *The Volta Review*, 64, 401-403.
- Canadian Hearing Society (1989). Press release, July 29, 1989.
- Carney, A.E., (1986). 'The role of tactile aids in rehabilitative audiology', in M.J. Collins et al. (editors), Sensorineural Hearing Loss: Mechanisms, Diagnosis and Treatment. Univ. of Iowa, Iowa City, Iowa.
- Carney, Arlene Earley (1988). "Vibrotactile Perception of Segmental Features of Speech: A Comparison of Single-Channel and Multichannel Instruments", *JSHR*, Vol. 31, 438-448.
- Carney, Arlene Earley, and Beachler, Cindy R. (1986). "Vibrotactile perception of suprasegmental features of speech: A comparison of single-channel and multichannel instruments", *JASA*, 79 (1), 131-140.
- Chan, Lester C.M. and Cheung, Y.S. (1986). "Analysis and Recognition of Isolated Putonghua Vowels by Karhunen-Loeve Transformation Techniques", *Speech Communication* 5, 299-330.
- Cole, R.A. and Scott, B.L. (1974). "Toward a theory of speech perception", *Psychological Review*, 81, 348-374.
- Colton, Raymond H. and Cooker, Harry S., (1968). "Perceived Nasality in the Speech of the Deaf", *JSHR*, 11, 553-559.

Cowan, R.S.C., Alcantara, J.I., Flamey, P.J. and Clark, G.M. (1988). "Preliminary evaluation of a multichannel electrotactile speech processor", JASA, 83(6), 2328-2338.

Craig, James C. (1985). "Tactile pattern perception and its perturbations", JASA, 77(1), 238-246.

Craig, James C. and Sherrick, Carl E. (1982). "Dynamic Tactile Displays", in Schiff and Foulke (eds.), Tactual Perception, A Source Book, Cambridge Univ. Press, 209-232.

De Filippo, Carol Lee (1984). "Laboratory Projects in Tactile Aids to Lipreading", Ear and Hearing, Vol. 5, No. 4, 211-227.

DeFilippo, C.L., and Scott, B.L. (1978). "A method for training and evaluating the reception of ongoing speech", JASA, 63, 1186-1192.

Delattre, Pierre C., Liberman, Alvin M., and Cooper, Franklin S. (1955). "Acoustic loci and transitional cues for consonants", JASA, 27, 769-773.

DFDP User's Manual (1986). Atlanta Signal Processors Inc.

Dudley, Hommer (1976). "The Carrier Nature of Speech", in Acoustic Phonetics, D.B. Fry, editor, Cambridge University Press, Cambridge, 21-30.

Eisenberg, L.S. et. al. (1983). "Cochlear Implants in Children", Ear and Hearing, 4, 41-50.

Electronic Design (1986). Special DSP Issue, February 20.

Engebretson, A. Maynard and O'Connell, Michael P. (1986). "Implementation of a Microprocessor-Based Tactile Hearing Prosthesis", IEEE Trans. on Biomedical Engineering, Vol. BME-33, No. 7, 712-716.

Engebretson, A. Maynard, Morley, Robert E., and O'Connell, Micheal P. (1986). "A Wearable, Pocket-Sized Processor for Digital Hearing Aid and Other Hearing Prostheses Applications", IEEE ICASSP 86 Tokyo, 625-628.

Erber, N.P., (1974). In R.E. Stark (Ed.), Sensory Capabilities of Hearing Impaired Children. Baltimore University Park Press, 69-74.

Erber, Norman P. (1979). "Speech Perception by Profoundly Hearing-Impaired Children", JSHD, XLIV, Aug., 255-270.

Erber, Norman, P. (1973). "Visual Perception of Speech by Deaf Children: Recent Developments and Continuing Needs.", JSHD, XXXIX, 2, 178-185.

Erickson, K.T. and Michel, A.N. (1985). "Stability Analysis of Fixed-Point Digital Filters Using Generated Lyapunov Functions--Parts I and II", IEEE Trans. Circuits and Systems, CAS-32, 113-142.

Fairbanks, G., (1940). "Recent experimental investigations of vocal pitch in speech", JASA, 11, 457-466.

Franklin, Barbara and Saunders, Frank (1984). "A Comparison Between a 32- and 16-Channel Electrotactile Device", presentation to ASA meeting, 1984.

Friel-Patti, Sandy and Roeser, Ross J. (1983). "Evaluating Changes in the Communication Skills of Deaf Children Using Vibrotactile Stimulation", Ear and Hearing, Vol. 4, No. 1, 31-40.

Frost, B.J. (1987). Personal Communication.

Frost, Barry (1988). Personal Communication.

Fry, D.B., Abramson, Arthur S., Eimas, Peter D., and Liberman, Alvin M., (1962). "The identification and discrimination of synthetic vowels", Language and Speech, vol. 5, 171-189.

Gault, R.H. (1924). "Progress in experimants on tactual interpretation of oral speech", J. Abnorm. Soc. Psych., 14, 155-159.

Ibid, (1934). "An interpretation of vibrotactile phenomena", JASA, 5, 252-254.

Gedeon, Ibrahim J. (1990). "LPC based speech coder for cochlear implant system", M.Eng. Thesis, Carleton University, Ottawa, Canada.

Geldard, F.A. and Sherrick, C.E. (1971). "Cutaneous temporal acuity", Princeton Cutaneous Research Project, 18.

Geldard, Frank A. (1960). "Some Neglected Possibilities of Communication", Science, Volume 131, Number 3413, 1583-1588.

General Instruments (1988). Zeta noise blocker information package.

Gescheider, G.A. (1973). "Temporal relations in cutaneous stimulation", in F.A. Geldard (ed.), Cutaneous communication systems and devices, Psychonomic Society, Austin TX., 33-37.

Gold, B. and Rader, C.M. (1969). Digital Processing of Signals, McGraw-Hill, New York.

Grant, Ken W., Ardell, LeeAnn H., Kuhl, Patricia K., and Sparks, David W. (1985). "The contribution of fundamental frequency, amplitude envelope, and voicing duration cues to speechreading in normal-hearing subjects", *JASA*, 77, (2), 671-677.

Guelke, R.W., and Huyssen, M.J. (1956). "Development of Apparatus for the Analysis of Sound by the Sense of Touch", *JASA*, 31, (6), 799-809.

Hack, Zarita Caplan, and Erber, Norman P. (1982). "Auditory, Visual, and Auditory-Visual Perception of Vowels by Hearing-Impaired Children", *JSHR*, 25, 100-107.

Hamsher, Donald H. (1967). Communication System Engineering Handbook, McGraw-Hill Book Company, New York.

Harman, H.H. (1976). Modern Factor Analysis, Chicago Univ. Press, Chicago.

Hawkes, G.R. (1961). "Cutaneous discrimination of electrical intensity", *Am. J. Psychol.*, 74, 45-53.

Hawkes, G.R. and Warm, J.S. (1960). "The sensory range of electrical stimulation of the skin", *Am. J. Psychol.*, 73, 485-487.

Hotelling, Harold (1933). "Analysis of a Complex of Statistical Variables Into Principal Components", *Jour. of Educational Psychol.*, 24, 417-441, and 498-520.

Houde, R.A., and Levitt, H. eds. (1980). Sensory Aids for the Hearing Impaired, IEEE.

IEEE ASSP Magazine (1986). Special DSP Issue, L.R. Morris (guest ed.) December 1986.

Ifukube, Tohru (1982). "A Cued Tactual Vocoder", in J. Raviv (ed.), Uses of Computers in Aiding the Disabled, North-Holland Publishing Co., 197-213.

Ifukube, Tohru (1989). *JASA*

Jackson, L.B. (1970). "Roundoff Noise Analysis for Fixed-Point Digital Filters Realized in Cascade or Parallel Form", *IEEE Trans. on Aud. & Electroac.*, Vol. AU-18, 107-122.

Kaiser, J.F. (1965). "Some Practical Considerations in the Realization of Linear Digital Filters", *Proceedings of the Third Allerton Conference on Circuit and System Theory*, 621-633.

Katz, Allen and Wentzel, Craig (1979). "A Portable Tactile Aid for the Deaf Using a 280 Microprocessor Controller",

Kenshalo, Dan R. (1978). "Biophysics and Psychophysics of Feeling", in Edward C. Carterette and Morton P. Friedman (eds.), *Handbook of Perception Volume VIB Feeling and Hurting*, Academic Press, New York, 29-74.

Kiang, N.Y.S. and Moxon, E.C. (1974). "Tails of tuning curves of auditory nerve fibers", *JASA*, 55, 620-630.

Kiang, N.Y.S., and Moxon, E.C. (1974). "Tails of tuning curves of auditory-nerve fibres", *JASA*, 55, 620-630.

Kirman, Jacob H., (1973). "Tactile Communication of Speech: A Review and an Analysis", *Psychol. Bull.*, vol. 80, 54-74.

Klatt, D.H. (1987). "Review of Text-to-Speech Conversion for English", *JASA*, 82, 737-793.

Klein, W., Plomp, R., and Pols, L.C.W. (1970). "Vowel Spectra, Vowel Spaces, and Vowel Identification", *JASA*, 48(4), 999-1009.

Kramer, H.P. and Mathews, M.V. (1956). "A Linear Coding for Transmitting a Set of Correlated Signals", *IRE Trans. Inform. Theory*, IT-2, 41-46.

Lennig, M. (1984). Development of Normalization of Speech for Automatic Recognition, Final Report: Gov. of Canada, Dept. of Supply and Services Contract 9SES1-00122, Bell-Northern Research. Montreal, Canada.

Leob, Gerald E. (1985). "Cochlear Implants in Children: Panel Discussion", Cochlear Implants, eds. Schindler, R.A. and Merzenich, M.M., Raven Press, N.Y., 575-587.

Levitt, Harry (1972). "Acoustic Analysis of Deaf Speech Using Digital Processing Techniques", *IEEE Trans. Audio and Electroacoustics*, March.

Levitt, Harry (1973). "Speech Processing Aids for the Deaf: An Overview", *I.E.E.E. Trans. on Audio & Electroacoustics*, vol. AU-21, June, 269-273.

Levitt, Harry, (1973). "Speech Processing Aids for the Deaf: An Overview", *IEEE Transactions on Audio and Electroacoustics*, Vol. AU-21, No. 3, 269-273.

Levitt, Harry, (1974). In R.E. Stark (Ed.), Sensory Capabilities of Hearing Impaired Children. Baltimore University Park Press, 73-76.

Li, K.-P., Hughs, G.W. and House, A.S. (1969). "Correlation Characteristics and Dimensionality of Speech Spectra", JASA, 46(4), 1019-1025.

Markides, Andrea* (1979). "Speechreading (Lipreading)", Child: Care, Health, and Development, 5, 93-101.

Marrin, Ken (1986). "Six DSP processors tackle high-end signal-processing applications", Computer Design, March 1, 21-25.

Mason, J.L., Scilley, P.L. and Frost, B.J. (1981). "A vibrotactile auditory prosthetic device", International Electrical Electronics Conference Digest, 104-106.

Matlin, Margaret W. (1988). Sensation and Perception, Allyn and Bacon, Inc., Toronto, pp 38-39.

McEntire, Robert H. (1972). "Information Rate via Vibrotactile Two Dimensional 'Phantom' Sensation", IEEE EMBS Conf. 1972, 342-345.

Meadow, K.P. (1980). Deafness and Child Development, Univ. of Calif. Press, L.A., 26-31.

Miller, M.A. (1968). "Speech and voice patterns associated with hearing impairment", Audicibel, 17, 162-167.

Morris, L.R. (1987). Personal Communication.

Morris, L.R. and Barszczewski, P. (1989). "Algorithms, Hardware, and Software for a Digital Signal Processor Microcomputer-Based Speech Processor in a Multielectrode Cochlear Implant System", IEEE Trans. on Biomed. Eng., Vol. 36 #6, June.

Naik, Jayant M. and Doddington, George R. (1986). "High Performance Speaker Verification Using Principal Spectral Components", IEEE ICASSP86, 881-884.

Nasa - -from scien. american

National Academy of Sciences (1971). "Auditory and Speech Impairments (A Pre-conference Primer)", reprinted from Sensory Training Aid for the Hearing Impaired.

Nickerson, R.S., Stevens, K.N., Boothroyd, A., and Rollins, A., (1974). "Some observations on timing in speech of the deaf and hearing speakers.", BBN Report No. 2905, September.

Nickerson, Raymond S. and Stevens, Kenneth N., (1973). "Teaching Speech to the Deaf: Can a Computer Help?", IEEE Trans. A & EA, vol. AU-21, 445-455.

Norber, E. Harris, (1967). "Articulation of the Deaf", *Exceptional Children*, 33, 611-621.

Norton, S.J., Schultz, M.C., Reed, C.M., Braida, L.D., Durlach, N.I., Rabinowitz, W.M. and Chomsky, C. (1977). "Analytical Study of the Tadoma Method: Background and Preliminary Results", *JSHR*, 20, 625-637.

Nucleus Inc. (1987). *Nucleus Cochlear Implant Brochure and Specifications*.

O'Connor, J.D., Gerstman, L.J., Liberman, A.M., Delattre, P.C., and Cooper, F.S., (1957). "Acoustic cues for the perception of initial /w,j,r,l/ in English", in *Acoustic Phonetics*, Fry, D.B., editor (1976), Cambridge University Press, Cambridge, 298-314.

Oppenheim, Alan V. and Schafer, Ronald W. (1975). *Digital Signal Processing*, Prentice-Hall Inc., Englewood Cliffs, New Jersey, 143-158.

Osberger, M.J. and Neely, S.J., (1985). "Acoustic Distinctions Between Intelligible and Unintelligible Vowels in the Speech of the Deaf", Paper presented to the 109th Annual Meeting of the ASA.

Papoulis, A. (1965). Probability, Random Variables and Stochastic Processes, McGraw Hill, New York.

Pardo, J.M. (1982). "Vocal Tract Shape Analysis for Children", *I.C.A.S.S.P.*, p763-766.

Parks, T.W. and Burrus, C.S. (1987). Digital Filter Design, John Wiley and Sons, New York.

Pickett, J.M. (1982). "Hearing through sight and feeling", *I.E.E.E. Spectrum*, April, 37-41.

Pickett, J.M. and McFarland, William (1985). "Auditory Implants and Tactile Aids for the Profoundly Deaf", *JSHR*, Vol. 28, 134-150.

Pickett, J.M., (1963). "Tactual Communication of Speech Sounds to the Deaf: Comparison with Lipreading", *JSHR*, vol. 28, 315-330.

Pickett, J.M., and McFarland, William, (1985). "Auditory Implants and Tactile Aids for the Profoundly Deaf", *JSHR*, Vol. 28, 134-150.

Pickett, J.M., and Pickett, B. Horenstein (1963). "Communication of Speech Sounds by a Tactual Vocoder", *JSHR*, 6, 207-222.

Plomp, R., Pols, L.C.W. and van de Geer, J.P. (1966). "Dimensional Analysis of Vowel Spectra", JASA, 41(3), 707-712.

Pols, Louis C.W. (1971). "Real-Time Recognition of Spoken Word", IEEE Trans. on Computers, Vol. C-20, No. 9, 972-978.

Porter, Robert J., et al. (1987). "Rate versus duration of format transition as a discriminative cue", JASA Sup. 1, 82, S81.

Povel, D-J and Wansink, M. (1986). "A Computer-controlled Vowel Corrector for the Hearing Impaired", J. of Speech and Hearing Research, Vol. 29, 99-105.

Proctor, Adele (1988). "Project: Tactile Aid", Information for prospective subjects and their parents, Northeastern University.

Proctor, Adele, and Bickley, Corine (1988). "Elicitation of Spontaneous Speech Production from Preschool Aged Tactile Aid Users", Presented at the Acoustical Society of America Meeting, May 16-20, Seattle WA.

Proctor, Adele, and Goldstein Jr., Moise H., (1987). "Tactile Aid Usage in Young Deaf Children: Case Studies", Submitted to JSHR.

Ibid, (1985). "Tactile aids for profoundly deaf children", JASA, 77, (1), 258-265.

Ibid, (1983). "Development of Lexical Comprehension in a Profoundly Deaf Child Using a Wearable, Vibrotactile Communication Aid", Language, Speech and Hearing Services in School, Vol. 14, No. 3, 138-149.

Rabiner, Lawrence R. and Gold, Bernard (1979). Theory and Applications of Digital Signal Processing, Prentice Hall, N.J., pp205-214.

Rajasekaran, Periyagaram K. and Doddington, George R. (1985). "Speech Recognition in the F-16 Cockpit Using Principal Spectral Components", IEEE ICASSP85, 882-885.

Reed, C.M., Rabinowitz, W.M., Durlach, N.I., Braida, L.D., Conway-Fithian, S. and Schultz, M.C. (1985). "Research on the Tadoma Method of Speech Communication", JASA, 77, 247-257.

Resnick, Steffi B., et al. (1986). "Report of the Ad Hoc Committee on Cochlear Implants", ASHA, Vol. 28, No. 4, 29-52.

Rothenberg, Martin and Molitor, Richard D. (1979). "Encoding voice fundamental frequency into vibrotactile frequency", JASA, 66(4), 1029-1038.

Rothenberg, Martin, Verrillo, Ronald T., Zahorian, Stephen A., Brachman, Michael L., and Bloanowski Jr., Stanley J. (1977). "Vibrotactile frequency for encoding a speech parameter", JASA, 62(4), 1003-1012.

Saunders, F.A. (1977). "Recommended Procedures for Electrocutaneous Displays", in F.T. Hambrecht and J.B. Reswick (eds.), Biomedical Engineering and Instrumentation, Vol. 3.

Saunders, Frank A. (1973). "An Electrotactile Sound Detector for the Deaf", IEEE Trans. Audio and Electroac., Vol. AU-21, No. 3, 285-287.

Saunders, Frank A., Hill, William A., and Simpson, Carol A. (1976). "Speech Perception Via the Tactile Mode: Progress Report", IEEE ICASSP, 594-597.

Schafer, Ronald W., and Rabiner, Lawrence R. (1975). "Digital Representations of Speech Signals", Proceedings of the IEEE, Vol. 63, 662-667.

Schein, Jerome D., and Delk, Marcus T. (1974). The Deaf Population of the United States, National Association of the Deaf, Maryland, USA.

Schulte, K. (1977). "Acquisition of speech by means of definable tactual and visual kinetic transfers". In J.M. Pickett (Ed.), Papers from the Research Conference on Speech-Processing Aids for the Deaf, Gallaudet College, Washington D.C., 52-55.

Scott, B.L. (1976). "Temporal factors in vowel perception", JASA, 60, 1354-1365.

Scott, Brian L. (1980). "Speech as Patterns in Time", in R.A. Cole (ed.), Perception and Production of Fluent Speech, Lawrence Erlbaum Assoc., Hillsdale N.J., 51-71.

Scott, Brian L. (1981). "Tactile Aid to Speech Reception", United States Patent #4,250,637.

Searle, C.L., Jacobson, J.Z. and Kimberly, B.P. (1978). "Speech as patterns in the 3-space of time and frequency", Proceedings of the 14th Annual Carnegie Symposium on Cognition, Pittsburg.

Searle, C.L., Jacobson, J.Z. and Rayment, S.G. (1979). "Stop consonant discrimination based on human audition", JASA, 65, (3), 799-809.

Searle, Campbell L. (1985). "Research on Speech and Music Perception", Unpublished tutorial manuscript, Research Laboratory of Electronics, Massachusetts Institute of Technology.

Sedra, Adel S., and Smith, Kenneth C. (1982). Microelectronic Circuits, Holt, Rinehart and Winston, New York; 608-614.

Sherrick, Carl E. (1984). "Basic and applied research on tactile aids for deaf people: Progress and prospects", JASA, 75, (5), 1325-1342.

Ibid, (1975). "The Art of Tactile Communication", American Psychologist, March, 353-360.

Shigenaga, M. and Hirotaka, K. (1986). "Speech Training For Handicapped Children Using Vocal Tract Lateral Shapes", I.C.A.S.S.P., 637-640.

Simmons, F. Blair (1985). "Cochlear Implants in Children: Panel Discussion", Cochlear Implants, eds. Schindler, R.A. and Merzenich, M.M., Raven Press, N.Y., 575-587.

Smith, Andre (1988). Head of Audiology. Children's Hospital of Eastern Ontario (CHEO), Ottawa, Ontario, Canada. Personal Communication.

Smith, C.M., (1973). "Residual hearing and speech production in deaf children." Communication Sciences Laboratory, Doctoral Program in Speech and Hearing Sciences, CUNY, April 15, 1973.

Sparks, David W., Ardell, Lee Ann, Bourgeois, Michelle, Wiedmer, Beatrix and Kuhl, Patricia (1982). "Investigating the MESA (Multipoint Electrode Speech Aid): The transmission of connected discourse", JASA, 65(3), 810-815.

Spens, K.E. (1980). "Tactile speech conveying aids for the deaf: A comparison", Proceedings of the 10th ICA, Sidney.

Spens, K.E. and Plant, G. (1983). "A tactual 'hearing' aid for the deaf", Speech Transmission Laboratory Quarterly Progress and Status Report, Stockholm, Sweden, 52-56.

Stark, R.E. (1971). "The Use of Real-Time Visual Displays of Speech in the Training of a Profoundly Deaf, Non-Speaking Child: A Case Report", J. Speech Hear. Disord. 36, 297-409.

Stevens, K.N., Nickerson, R.S., Boothroyd, A., and Rollins, A., (1974). "Assesment of nasality in the speech of deaf children", BBN Report No. 2902, September.

Stevens, Kenneth N. and Blumstein, Sheila E. (1981). "The Search for Invariant Acoustic Correlates of Phonetic Features", in Peter D. Eimas and Joanne L. Miller (editors), Perspectives on the Study of Speech, Lawrence Erlbaum Associates, Publishers, Hillsdale, New Jersey, 1-38.

Stevens, Stanley Smith and Davis, Hallowell (1983). Hearing, Its Psychology and Physiology, American Institute of Physics for the Acoustical Society of America, pp248-287.

Strong, W.J. (1975). "Speech aids for the profoundly/severely hearing impaired: Requirements, overview, and projections", *Volta Review*, 77, 536-556.

SWDS User's Guide (1988). Texas Instruments Inc.

Szeto, Andrw Y.J. and Saunders, Frank A. (1982). "Electrocutaneous Stimulation for Sensory Communication in Rehabilitation Engineering", *IEEE Trans. Biom. Eng.*, Vol. BME-29, No. 4, 300-308.

Tachi, Susumu, Tanie, Kazuo, Komoriya, Kiyoshi and Abe, Minoru (1985). "Electrocutaneous Communication in a Guide Dog Robot (MELDOG)", *IEEE Trans. on Biomed. Eng.*, Vol. BME-32, No. 7, 461-469.

Tachi, Susumu, Tanie, Kazuo, Komoriya, Kiyoshi, Asaba, Kiichiro, Tomita, Yoshiyuki and Abe, Minoru (1980). "Information Transmission by Two-Dimensional Electrocutaneous Phantom Sensation", *J. of Jap. Mech. Eng. Laboratory*.

Texas Instruments (1988). "Details on Signal Processing", Issue 17.

Texas Instruments (1988). Second Generation User's Guide.

Texas Instruments (1989). "FYI Bulletin", Newsletter on new product announcements, March.

Tyler, Richard S., et al. (1986). "Cochlear Implant Program at the University of Iowa", in M.J. Collins et al. (editors), Sensorineural Hearing Loss: Mechanisms, Diagnosis and Treatment. Univ. of Iowa, Iowa City, Iowa.

United States Patent #4,728,934, (1988). "Tactile Stimulation Device for Hearing-Impaired Individuals", Assignee: Siemens Aktiengesellschaft, Berlin FRG.

Upton, H.W. (1968). "Wearable eye glass speechreading aid", *Amer. Annals of the Deaf*, 113, 222-229.

Verrillo, Ronald T. (1970). "Psychophysics of vibrotactile stimulation", *JASA*, 43 (2).

Verrillo, Ronald T. (1970). "Subjective Magnitude Functions for Vibrotaction", IEEE Trans. on Man-Machine Sys., Vol. MMS-11, No. 1, 19-24.

von Békésy, G. (1957). "Neural Volleys and the Similarity Between Some Sensations Produced by Tones and by Skin Vibrations", JASA, 29, 1059-1060.

von Békésy, G. (1959). "Synchronism of Neural Discharges and Their Demultiplication in Pitch Perception on the Skin and in Hearing", JASA, 31, 338-349.

Walden, Brian E., Prosek, Robert A., Montgomery, Allen A., Scherr, Charlene K., and Jones, Carla J (1977). "Effects of Trainging on the Visual Recognition of Consonants", JSHR, 20, 130-145.

Walsh, Margaret A., and Diehl, Randy L. (1987). "Auditory factors in the perception of stops and glides", JASA Supl 1, 82, S80.

Weisenberger, Janet M. (1986). "Measures of Temporal Processing of Vibrotactile Stimuli", presented at the ASA meeting, Cleveland, OH., May.

Weisenberger, Janet M. (1988a). "Evaluation of Tactile Aids for the Hearing-Impaired", ICAART, Montreal, 224-223.

Weisenberger, Janet M. (1988b). "Effects of number of channels on speech perception with tactile aids", presentation to the 116th meeting of ASA, Honolulu, Hawaii.

Weisenberger, Janet M. and Russell, Ann F. (1986). "A Comparison of Two Single-Channel Vibrotactile Aids for the Deaf", Presentation to the Acoustical Society of America, Annual Meeting, Dec. 1986.

Weisenberger, Janet M., and Miller, James D. (1987). "The role of tactile aids in providing information about acoustic stimuli", JASA, 82 (3), 906-916.

Witten, I.H. (1982). Principles of Computer Speech, Academic Press, Toronto, 63-67.

Yeni-Komshian, Grace H., and Goldstein, Moise H. (1977). "Identification of speech sounds displayed on a vibrotactile vocoder", JASA, 62(1), 194-198.

Yilmaz, H. (1967). "A theory of speech perception", Bulletin of Mathematical Biophysics, 29, 793-824.

Yilmaz, H. (1968). "A theory of speech perception, II", Bulletin of Mathematical Biophysics, 30, 455-479.

Zahorian, Stephen A. (1988a). "The application of principal-components spectral analysis to sensory aids for the hearing impaired", submitted for publication.

Zahorian, Stephen A. (1988b). Personal Communication.

Zahorian, Stephen A. and Rothenberg, Martin (1981). "Principal-components analysis for low-redundancy encoding of speech spectra", JASA, 69(3), 832-845.

Zahorian, Stephen A., and Jagharghi, Amir J. (1985). "Color Speech Display as a Speech Training Aid for the Deaf", paper presented at the 109th Meeting of the ASA, Austin TX., April.

Zimmerman, Gerald, and Rettaliata, P., (1981). "Articulatory Patterns of an Adventitiously Deaf Speaker", JSHR, (24), 169-178.

Bibliography

VISUAL AIDS

Goldberg, Aaron J (1972). "A Visual Feature Indicator for the Severely Hard of Hearing", IEEE Tran. A&E, vol. **AU-20**, no. 1, 16-23.

Goldstein, M.H. and Stark, R.E. (1976). "Modification of vocalizations of preschool deaf children by vibrotactile and visual displays", JASA, vol. **59**, 1477-1481.

House, A.S. et Al. (1967). "Perception of Visual Transforms of Speech Stimuli: Learning Simple Syllables", Conference on Speech Analyzing Aids for the Deaf, Gallaudet College, Washington D.C., 215-221.

Montgomery, Allen A. (1984). "Training Auditory-Visual Speech Reception in Adults with Moderate Sensorineural Hearing Loss", Ear and Hearing, Vol. 5, No. 1, 30-36.

Pickett, J.M. and Constam, Alfred (1972). "A Visual Speech Trainer with Simplified Indication of Vowel Spectrum", ieee book, 253-258.

Pronovost, Wilbert et Al. (1968). "The Voice Visualizer", ieee book, 230-238.

Stark, Rachel E. et Al. (1970). "Preliminary Work with the New Bell Telephone Visible Speech Translator", IEEE, 205-214.

Walden, Brian E. et Al. (1977). "Effects of Training on the Visual Recognition of Corsonants", JSHR, Vol. 20, No. 1, 130-145.

Walden, Brian E. et Al. (1981). "Some Effects of Training on Speech Recognition by Hearing-Impaired Adults", JSHR, **24**, 207-216.

COMPUTER-AIDED TRAINING

Boothroyd, Arthur and Levitt, Harry (1982). "Computers and the Education of Hearing-Impaired Children: Possibilities and Limitations", Uses of Computers in Aiding the Disabled, North-Holland Publishing Co., 285-293.

Di Benedette, Maria Domenica et Al. (1982). "Phonetic Recognition to Assist Lip-Reading for Deaf Children", ICASSP, 739-742.

Haton, Marie-Christine and Haton, Jean-Paul (1982). "Computer-Aided Speech Analysis and Training for the Hearing Impaired", Uses of Computers in Aiding the Disabled, North-Holland Publishing, 185-195.

Nickerson, Raymond S. and Stevens, Kenneth N. (1973). "Teaching Speech to the Deaf: Can a Computer Help?", IEEE Trans. A & EA, vol. AU-21, 445-455.

COCHLEAR IMPLANTS

Canada:

Harrison, R.V., et Al. (1987). "Technical development fo an implantable cochlear prosthesis in Canada", J. of Otolaryngology, 16:5, 311-315.

Misc.:

Genin, J. and Charachon, R. (1986). "Sound Processing for Cochlear Implant", ICASSP, 621-623.

Ifukube, Tohru and White, Robert L. (1986). "A Speech Processor with Lateral Inhibition for an Eight Channel Cochlea Implant and its Evaluation by Subject Testing", ICASSP, 617-620.

Geisler, C. Daniel (1987). "Coding of Acoustic Signals on the Auditory Nerve", IEEE Eng. in Med. and Bio. Mag., June, 22-28.

Tyler, Richard S., et al. (1984). "Cochlear Implant Program at the University of Iowa", in T.J. Glatcke (ed.), Proceedings of the Scott Reger Conference, University Park Press, Baltimore.

Feigenbaum, Ernest (1987). "Cochlear Implant Devices for the Profoundly Hearing Impaired", IEEE End. in Med. and Bio. Mag., June, 10-21.

Resnick, Steffi B., et al. (1986). "Report of the Ad Hoc Committee on Cochlear Implants", *Asha*, Vol. 28, No. 4, 29-53.

Tyler, Richard S., et al. (1987). "Vowel and Consonant Confusions Among Cochlear Implant Patients: Do Different Implants Make a Difference?", *Annals of Otology, Rhinology, & Laryngology*, Vol. 96, No. 1, part 2, 141-144.

Implants in Children:

Tyler, Richard S., et al. (1986). "Clinical Objectives and Research-Design Issues for Cochlear Implants in Children", *Seminars in Hearing*, Vol. 7, No. 4, 433-440.

Tyler, Richard S., Davis, Julia M., and Lansing, Charissa R. (1987). "Cochlear Implants in Young Children", *ASHA*, April, 41-49.

Australia - Nucleus:

Millar, J.B., et Al. (1984). "Speech Processing for Cochlear Implant Prosthesis", *ASLHA*, vol. 27, 280-296.

Blamey, P.J., et Al. (1985). "A comparison of three speech coding strategies using an acoustic model of a cochlear implant", *JASA*, 77(1), 209-217.

Cochlear Corp., A Sound Improvement, brochure, Cochlear Corp., 61 Inverness Dr. E., Englewood, CO., U.S.A.

Cochlear Corp., Nucleus Mini 22 Cochlear Implant System, brochure, Cochlear Corp., 61 Inverness Dr. E., Englewood, CO., U.S.A.

Tyler, Richard S., et al. (1984). "Initial Iowa Results with the Multichannel Cochlear Implant from Melbourne", *JSHR*, Vol. 27, 596-604.

ibid, (1986). "Previous Experience as a Confounding Factor in Comparing Cochlear-Implant Processing Schemes", *JSHR*, Vol. 29, 282-287.

Stanford:

White, Robert L., et Al. (1982). "The Use of Computers in the Development of a Cochlear Prosthesis", Uses of Computers in Aiding the Disabled, ed. Raviv, E., North-Holland Publishing, 147-165.

White, Robert L. (1987). "System Design of a Cochlear Implant", *IEEE Eng. in Med. and Bio. Mag.*, June, 42-46.

Atlas, L.E., et Al. (1983). "Results of stimulus and speech coding schemes applied to multichannel electrodes", Annals of the New York Academy of Sciences, 405, 377-386.

3M, House, Vienna:

Hochmair, E.S., and Hochmair-Desoyer, I.J. (1983). "Percepts elicited by different speech coding strategies", Annals of the New York Academy of Sciences, 405, 268-279.

Hochmair-Desoyer, I.J., et Al. (1981). "Four years experience with cochlear prostheses", Medical Progress Through Technology, 8, 107-119.

ibid, (1980). "Cochlear prostheses in use: Recent speech comprehension results", Archives of Otorhinolaryngology, 229, 81-98.

3M Corp. , 3M Vienna Cochlear Implant System: Questions & Answers, brochure, 3M Vienna Cochlear Implant Program, Otologic Products/3M, 225-5N-03, 3M Center, St. Paul MN, U.S.A.

3M Corp., Vienna Cochlear Implant, Hear the World, Hearing Health Care/3M, Building 225-5n-03, #M Center, St. Paul, MN, U.S.A.

Punch, Jerry, L., et al. (1987). "Relationships among Selected Measures of Single-Channel Cochlear Implant Performance", Ear nad Hearing, Vol. 8, No. 1, 37-43.

Paris, Bertin:

Chouard, C.H., et Al. (1983). "Long-term results of the multichannel cochlear implant", Annals of the New York Academy of Sciences, 405, 387-411.

Chouard, C.H. (1980). "Surgical rehabilitation of total deafness with the multichannel cochlear implant", Audiology, 19, 137-145.

ibid, (1978). "Multiple intracochlear electrodes for rehabilitation of total deafness", in M.E. Glasscock (Ed.), Otolaryngological clinics of North America (vol. 11), Philadelphia: W.B. Saunders.

Bertin et Compagnie, The Bertin Cochlear Implants, brochure, Bertin et Compagnie.

Symbion, M.I.T.:

Dorman, Michael F., et al. (1988). "Auditory/phonetic categorization with the Symbion multichannel cochlear implant", *JASA*, 84(2), 501-510.

Eddington, D.K. (1983). "Speech recognition in deaf subjects with multichannel intracochlear electrodes", in Cochlear Prostheses: An International Symposium, edited by C. Parkins and S. Anderson (NYAS, New York).

ibid, (1980). "Speech discrimination in deaf subjects with cochlear implants", *JASA*, 51, 885,891.

Research Triangle Institute:

Wilson, Blake S., et al. (1987). "Comparisons of Processing Strategies for Multichannel Auditory Prostheses", *IEEE Ninth Ann. Conf. Eng. in Med. and Biol. Soc.*, 1908-1910.

TACTILE AIDS

Overview:

"Aids for the Hearing Impaired", Report of the Panel on Communicative Disorders to the National Advisory Neurological and Communicative Disorders and Stroke Council, in Levitt J. (editor), IEEE Sensory Aids, 547-557.

"Tactile communication for the deaf", *JASA*, 76(3), 976-977.

Eilers, Rebecca E., Oller, D.K. and Lynch, Michael P. (1988). "The Development of Artificial Hearing Through the Skin", *ICAART 88- Montreal*, 220-221.

Houde, R.A. (1981). "Tactual and Visual Aids: Speech-Feature Displays", in Levitt J. (editor), IEEE Sensory Aids, 319-321.

Houde, R.A. (1981). "Visual and Tactual Aids for Speech Reception: Spectrum and Temporal Displays", in Levitt J. (editor), IEEE Sensory Aids, 243-246.

Kaplan, H. and Fernandes, C.C. (1986). "Assistive Devices for Deaf and Hard of Hearing People", *RESNA 9th Annual Conference, Minneapolis*, 65-68.

Kringlebotn, M. (1968). "Experiments with some Visual and Vibrotactile Aids for the Deaf", *Amer. Annals of the Deaf*, 113, 311-317.

Martin, M.C. (1985). "Alternatives to Cochlear Implants", in R.A. Schindler and M.M. Merzenich (eds.), Cochlear Implants, Raven Press, N.Y., 549-561.

Martony, Janos and Spens, Karl-Erik (1972). "Tactual and

Visual Lipreading Aids", 1972 Conf. Speech Comm. & Processing, 242-245.

Miller, J.O., (moderator) (1982). "Panel Discussion", in J. Raviv (ed.), Uses of Computers in Aiding the Disabled, North-Holland Publishing Co., 309-317.

Pickett, J.M. (1982). "Hearing through sight and feeling", IEEE Spectrum, April, 37-41.

Pickett, J.M. and Trybus, Raymond J. (1983). "The Gallaudet Rehabilitation Engineering Center for the Deaf and Hearing Impaired", 6th Annual Conference on Rehabilitation Engineering, San Diego, 318-320.

Risberg, A. (1969). "Critical Review Work on Speech Analyzing Hearing Aids", IEEE Trans. Audio. Electroacoust., vol. AU-17, 290-297.

Single Feature Devices:

Collins, M. Jane and Hurtig, Richard R. (1985). "Categorical Perception of Speech Sounds Via the Tactile Mode", JSHR, Volume 28, 594-598.

Spens, Karl-Erik and Martony, Janos (1972). "The Transposer as a Lipreading Aid", 1972 Conf. Speech Comm. & Processing, 246-248.

Summers, I.R. (1982). "An Improved Tactile Aid for the Deaf", IEE #258.

Szeto, Andrew Y. and Farrenkopf, Glenn (1988). "Optimization of Single Electrode Tactile Codes", ICAART 88 - Montreal, 222-223.

Walker, J.R.M., Fenn, G. and Smith, B.Z.D. (1987). "Soft-Talker: A Sound Level Monitor for the Hard-of-Hearing Using an Improved Tactile Transducer", J. Biomed. Eng., vol. 9, 177-179.

Walliker, John, Rosen, Stuart, and Fourcin, Adrian (1982). "Speech Pattern Prostheses for the Profoundly and Totally Deaf", IEE #258, 194-199.

Multifeature Tactile Aids:

Chan, Hoi-Young and O'Brien, Edward M. (1983). "Visual Perception of a Phonemic Code for a Tactile Hearing Aid", IEEE Conf. of the Eng. in Med. and Biol. Soc., 376-379.

Chan, Hoi-Young and O'Brien, Edward M. (1984). "Tactile Perception of a Phonemic Code for a Tactile Hearing Aid", IEEE Conf. of the Eng. in Med. and Biol. Soc., 375-378.
Green, B.G., Graig, J.C., Wilson, A.M., Pisoni, D.B. and

Rhodes, R.F. (1983). "Vibrotactile Identification of Vowels", JASA, 73, 1766-1778.

Kanevsky, Dmitry and Skurkovich, Grigori (1982). "A Vibrotactile Articulator", in J. Raviv (ed.), Uses of Computers in Aiding the Disabled, North-Holland Publishing Co., 273-278.

Liu, Ying (1987). "Tactile Hearing Prosthesis Based on Computer Speech Automatic Recognition", IEEE/Ninth Ann. Conf. Eng. in Med. and Biol. Soc., 1899-1900.

Miller, James D., Engebretson, A. Maynard, and De Filippo, Carol L. (1974). "Preliminary Research with a Three-Channel Vibrotactile Speech-Reception Aid for the Deaf", Gunnar Fant (ed.), Speech Communication, Vol. 4, John Wiley & Sons, N.Y., 341-347.

Ozdamar, O., Delgado, R.E. and Lopez, C. (1987). "FFT-Based Digital Vocoder for Artificial Hearing", IEEE/Ninth Ann. Conf. Eng. in Med. and Biol. Soc., 1892-1893.

Saunders, Frank A. (1973). "An Electrotactile Sound Detector for the Deaf", IEEE Trans. on Audio and Electroacou., Vol. AU-21, No. 3, 285-287.

Ueda, Yuichi and Watanabe, Akira (1986). "Assessment of Electronic Aids for the Hearing Impaired which Transmit Visible or Tactile Speech", IEEE ICASSP 86, Tokyo, 629-632.

Ueda, Yuichi and Watanabe, Akira (1987). "Visible/tactile vowel information to be transmitted to the hearing impaired", J.Acoust. Soc. Jpn. (E), 8, no. 3, 99-108.

END

03.03.92

FIN