

Final Report

January 1, 1996, through December 31, 1998

Speech Processors for Auditory Prostheses

NIH Contract N01-DC-6-2100

submitted by

Donald K. Eddington

Nico Garcia

Victor Noel

Joseph Tierney

Margaret Whearty

Massachusetts Institute of Technology

Research Laboratory of Electronics

Cambridge, MA

Summary of Salient Results

A new six-electrode current stimulation circuit, D/A circuit, microphone input conditioning circuit, and power conditioning circuit for use with the Geneva Wearable Processor (GWP) has been designed. A High Rate Stimulation System (HRSS), software controlled, provides arbitrary waveform samples with a 2 μ s smallest interval for implementing speech processors and psychophysical testing in the laboratory. New software allows pairwise processor comparisons in every-day listening conditions with the GWP. Continuous Interleaved Sampling Processor (CIS) implementations and fitting procedures are compared for three different research sites. Control capability for the Clarion Implanted Receiver exists in the laboratory.

Loudness growth studies for implanted subjects demonstrate CIS improvements when peak output currents from mapping functions are balanced for loudness and the functions emulate loudness growth for normal hearing subjects. We compare CIS longitudinal data for implanted subjects at MIT with subjects at RTI and UoG. Data from one bilaterally implanted subject demonstrates sensitivity to interaural time differences for cochleotopically matched electrodes, and no sensitivity when electrodes are mismatched. Finally we present speech reception data that motivated a processor design which provided more information to electrodes #1 and 2. The form of this information did not enhance speech reception performance.

Table of Contents

1	Introduction	2
2	Hardware Studies.....	2
3	Software Studies	6
4	Processor Studies	9
5	Bilateral Studies.....	25
6	Acoustic Simulations and New Processor Designs.....	32

1 Introduction

Work performed with the support of this contract was directed at the design, development, and evaluation of speech processors for use with auditory prostheses implanted in deaf humans. Major research efforts occurred in four areas: (1) developing and maintaining a laboratory based, software controlled, real time, speech processing facility where processor/stimulator algorithms for monaural and binaural eight-channel implants can be implemented/tested and a wide range of psychophysical measurements can be made, (2) using the laboratory facility to refine the sound processing algorithms used in the current commercial and laboratory processors, (3) using the laboratory facility to explore new sound processing algorithms for implanted subjects, and (4) designing and fabricating programmable, wearable speech processors/stimulators and using these systems to: (a) field test processor algorithms developed and tested in the laboratory, (b) evaluate the effects of learning using longitudinal evaluations of speech reception, and (c) compare asymptotic performance of different speech processors across subjects. This Final Report summarizes the progress associated all four areas.

2 Hardware Studies

2.1 Wearable Processors

The new wearable sound processing system, used as the basis for all wearable processor experiments under the present contract, was designed under the previous contract N01-DC-2-2402. The design and implementation were done in collaboration with the Microprocessor Laboratory of the Geneva Engineering School (GES), the University of Geneva (UoG), and the research Triangle Institute (RTI). Support for the work came from the NIH contract, a parallel contract at RTI, and from the Wilsdorf Foundation of Geneva. This Geneva Wearable Processor (GWP) replaces all of the external components of the Ineraid hardware system in approximately twenty subjects, and gives us the ability to implement a wide range of sound processors designed for moderate rate (<3,000 pulses/sec/channel), interleaved, monolateral stimulation.

Components replaced include: the sound processor, processor/earhook cable and connectors, earhook assembly, and the earhook/pedestal cable and connector as shown in Figure 1. The power sources for these processors consist of a four-cell battery pack that uses type 150AFH (roughly AA size) NimH rechargeable units. The charging devices, designed by the Engineering School, are also provided to our subjects ¹.

2.2 Improved Wearable Processors

In order to move from our current processor to one capable of testing processors based on high-rate, interleaved or simultaneous stimulation, several modifications were required. First the bandwidth of the current sources was increased to transmit the narrower pulse widths. At the same time, the power dissipation of these circuits as well as their size was reduced so that six of these current sources can be packaged in the existing GWP case. The following describes the design and testing work conducted with support of this contract that will lead to the production of the new processor units.

2.2.1 Stimulation System Design

The stimulation system assembly includes six D/A converters (all on a single chip) that are driven by serial input from the signal processing electronics. Each D/A output drives a current source (V/I) that produces corresponding output stimulation currents to be applied to implanted cochlear electrodes. This assembly also includes a battery voltage conditioning circuit that produces operating voltages for all of the analog and digital electronics. This assembly of V/I converters, D/A converters, and power conditioning circuits is packaged as a compact multi-chip module (MCM) that is integrated with the analog input circuits of the GWP. The V/I converters are capable of producing pulsatile stimuli with full-scale rise/fall times around 1 μ s, maximum output currents of ± 9 ma, and a minimum compliance of ± 10 volts.

Figure 2. shows a schematic of the latest design for the Draper Laboratory, V/I circuit (as the diagram indicates, two current sources are included in the design layout) ². This version of the V/I converter provides a number of advantages over the Howland circuit we have used in the past: better temperature stability, improved power supply rejection ratio, and a reduction in quiescent power dissipation.

2.2.2 Design of the Output Subsystem

In addition to the design and testing of the V/I circuit, mechanical/electrical design of the miniaturized output subsystem described above has also been completed. Based on studies of alternative packaging formats for size and mechanical stability, the choice was made to use a set of multi-chip modules (MCM) to be integrated on a common substrate. This format uses unpackaged IC die and miniaturized passive

elements. A design of the stimulator circuit (V/I) MCM has been completed that includes two V/I converters. An MCM design of the eight channel D/A unit has also been completed and is based upon the Analog Devices AD7568 Octal 12 Bit DAC. The D/A unit includes circuits that enable the sound processor DSP chip to prevent current stimulation until the entire portable processor system is operational and all internal tests are completed successfully. At the present time a power conditioning MCM layout has also been completed that is based upon a single DC to AC conversion oscillator circuit that is transformer coupled to provide isolated outputs that are rectified and smoothed to provide the required voltages.

Figure 3. presents the MCM subsystem layout. The figure shows each of the MCM units arrayed on the final substrate. The upper left hand unit is the voltage conditioning circuit which includes a small toroidal transformer and a voltage converter IC (the MAX 763) used to generate the oscillatory waveform which drives the transformer outputs. The three units arrayed along the top next to the voltage conditioning circuit are the MCM stimulating units (two V/I circuits each). Finally, the circuit spread out along the bottom margins of the substrate is the D/A unit. The entire substrate assembly is roughly 4.5 x 8.5 cm and occupies one side of one of the two boards that comprise the present GWP.

At the same time the wearable processor output circuits were undergoing redesign, we also updated and improved the analog input circuits. This resulted in a complete I/O design that is more flexible than in the present wearable processor and uses lower power technology.

2.2.3 New Analog Input Circuit Design

The present GWP analog input circuits that provide a digitized speech sample to the digital signal processing chip (DSP) use an A/D converter design (Burr Brown – ADS7807) that requires an external filter before the A/D stage to prevent aliasing. This external filter is set to eliminate frequencies above 6800 Hz in the present system. If a higher sampling rate than is presently used by the DSP (approximately 16 kHz) is desired to process a wider spectrum of analog signals, the cut-off frequency of the present filter needs to be increased. In the redesigned circuit, a new A/D converter part is used (Analog Devices-AD73311) that provides internal, flexible anti-aliasing filtering so that sampling frequencies and spectrum cut-off frequencies are linked as the sampling frequency is changed under program control. Besides the flexibility, such an arrangement also eliminates the space and power required by the antialiasing filter. Additionally, this new part provides several on-chip clock dividers that eliminate the need for several digital circuits used by the GWP. Taken together, this new design (shown in Figure 4.) provides more flexibility, more space and power efficiency, and improves the reliability of the overall system.

2.3 Laboratory Facilities

2.3.1 The High Rate Stimulation System (HRSS)

Since the beginning of the presently completed, speech processors contract, we have continued to upgrade our laboratory processing system so as to perform each of the research tasks required under the contract. Under test, as the present contract ended, is our DSP-based high rate stimulation system (HRSS) as shown in the upper right hand corner of Figure 5. The box labeled "percutaneous output" represents the segment of our laboratory sound processing system with the present capability of dual, eight-channel, 16-bit, D/A outputs that are used to drive percutaneously connected electrodes (through isolated current source circuits not shown). The system we have designed and tested is represented by the other two boxes of the "percutaneous output" system (arbitrary waveform generator and V/I converter). The DSP-based arbitrary waveform generator is controlled by a program that is downloaded from the SUN workstation and is able to produce arbitrary waveform samples with a 2 μ s sample interval. The generated waveforms can be amplitude (or time-shift) modulated in real-time by inputs from the dual floating-point DSP system. The arbitrary waveform generator off-loads the modulation task from the floating-point DSP system and allows additional processing power to be available for more complex speech processor implementations

Figure 6. presents a detailed view of the new system³. As shown, each channel includes a Motorola DSP 56002 running a program that is downloaded from the SUN workstation. This program then accepts real-time inputs from the floating-point Texas Instruments TMS320C30-based laboratory DSP system. The serial, digital data stream from each channel's 56002 that represent the 2 μ s waveform samples are optically isolated before being connected to the input of a 16-bit, 2 μ s D/A converter. At the present time we have completed the design, fabrication, and debugging of the DSP portion of the HRSS which is implemented on a small, 3"x4" printed circuit board. The opto-isolator circuits and D/A converter (Burr-Brown PCM56) have been integrated and debugged, and a new current source designed by the Draper Laboratory has been debugged and tested. This new current source circuit trades off additional power dissipation (not a critical concern in the laboratory setting) to make a simpler and less expensive circuit than the one used for the new wearable processor design and also includes a shield driver circuit.

3 Software Studies

3.1 Software Support for Wearable Processors

3.1.1 Converting Laboratory Processor CIS for GWP Use

The software mechanisms for converting laboratory processor code -- written to run on the floating point facility -- into the code necessary for the fixed-point processor used in the GWP were in place at the end of our earlier NIH contract. In order to produce real-time code for the CIS implementation on the GWP work was performed to integrate the I/O drivers and to insure safe startup and shutdown procedures.

The I/O driver was modified to control the 16-bit A/D converter used in the GWP and transmit data from the earpiece microphone to the DSP. At the same time, the driver was also modified to accommodate communication between the DSP and the D/A converter that drives the single current source circuit in the GWP; as well as controlling the multiplexing of this single source so as to drive one of six of the Ineraid electrodes. Software control of the multiplexed current source also included a startup procedure that allowed for the processor to begin from a reset condition and initialize all the channel stimulation output parameters before the user's electrodes were connected to the current source. During continuous use of the GWP, power supply monitoring guarantees that the processor shuts down in an equally graceful and safe manner when battery voltage is below a threshold (a slightly higher threshold allows for a preliminary warning to the user by flashing a light), because of battery drain or because the on-off switch has been used.

3.1.2 Modifying the CIS Loader to Load Two Mapping Function and I_{max} Sets

It became apparent during the course of the present contract that subjective evaluation and comparison of parameter modifications to our CIS implementation on the GWP would be facilitated if subjects could take home more than one processor implemented on the same GWP. Software has been developed that allows a CIS implementation to use one of two sets of downloaded mapping functions and the corresponding peak stimulation currents associated with each set of mapping functions. When the GWP is turned on, a user switch setting is sensed that directs the CIS software to one of the stored sets. In this way we can evaluate, for example, the use of logarithmic mapping functions versus mapping functions that reproduce normal hearing's loudness growth. Similar comparison/evaluations can be implemented depending upon the amount of additional control and memory required.

3.2 Software Support – Laboratory Facilities

3.2.1 Software Development for the HRSS

Substantial work has been completed on the software required to control the new stimulation system that is capable of producing arbitrary sampled waveforms with intervals as short as two microseconds. The new system replaces a hardware, state-machine based sequencer, with a bank of Motorola 56002 DSP processors. While the hardware sequencer was very fast, the new DSP software-based approach will allow for more flexibility in production of stimulation signals. Software was completed that allows the new system to perform its first intended task: psychophysical testing using waveforms of arbitrary shape (e.g., charge balanced tri-phasic pulses). The software written includes a module that allows experimenters to specify waveform shapes in a simple text file format as shown in Figure 6. This text file, generated from any word processor, is then converted and downloaded to one of the 56002 modules. The 56002 assembly language software that accepts the modulating signal from the floating-point, C30 DSP, modulates the waveform defined by the text file and outputs modulated waveform, is also completed. This software allows the floating-point, C30 DSP system to produce multichannel, real-time processor envelopes each of which can be used to amplitude modulate the arbitrary waveform set up in each channel of the HRSS. Since each 56002 of the HRSS can be programmed to generate a constant value rather than an arbitrary waveform, each 56002 can be used as a "pass through" to the isolated current sources so that waveform samples from processors implemented in the floating-point, C30 DSP system can be output from the isolated current outputs. This capability allows direct waveform inputs (albeit sampled) to the isolated V/I circuits. This means that a variety of hybrid processors can be implemented where some C30 DSP channels produce waveforms that are "passed through" to the V/I circuits and others are used to modulate the arbitrary stimulation waveforms specified for other channels of the HRSS.

Software development will be an ongoing effort as new research goes forward, and new stimulation waveforms are required. Indeed, this is a feature of the new DSP-based approach. Software completed at this time enables the new system to be used for all present psychophysical measurement work, as well as new sound processor simulations.

3.2.2 Control of Clarion Implants

Work has also been completed allowing control of implanted current stimulators manufactured by Advanced Bionics and to synchronize those stimuli with stimuli we deliver to Ineraid implants from our percutaneous system. While our present system has enabled us to test some new sound processing systems and conduct monolateral and bilateral psychophysical testing with time resolution of better than 1 μ s. Timing uncertainty between the bilateral stimuli presented to the left and right cochlear implant electrodes is of the order of ± 100 nanosec. We will discuss use of this facility in a later section on bilateral implants.

We received hardware, documentation and source code, developed at Advanced Bionics and at the House Ear Institute, that will, with some modifications and additions, provide the control we will require for experiments we plan to conduct in the future. Instead of performing a speech processing strategy, the Clarion sound processor will accept commands from an external serial port, translate them, and essentially pass them to the implanted receiver-stimulator device. These commands will allow complete control of the implant, including electrode configuration and stimulation current output for all 8 channels, at the full 76.9 us update rate. The investigator will be able to feed commands to the speech processor's serial port from a DSP processor, which in our case will be one of our laboratory system's TMS320C30s. The received source code has been shared with our engineering partners at Draper Laboratory, who will assist us in development of a software library for the C30 that will allow investigators to have high level access to the implant functions. Understanding the details of controlling the Clarion implant will then be used later, in our design of a research wearable system capable of controlling this same Clarion implant.

4 Processor Studies

4.1 Comparisons of CIS Structure and Fitting Procedures -- MIT, RTI, UoG

General CIS channel operations are shown in Figure 7. To produce current stimulation for six electrodes, six channels of such processing are implemented. As shown, front-end operations like automatic gain control (AGC), and/or a frequency dependent gain (preemphasis) operate between the microphone (MIC) and the set of band-pass filters that span the frequency range to be represented. Each channel consists of: (1) a band-pass filter; (2) an operation that measures the envelope variations in each band-pass filter's output waveform; (3) low-pass filtering of the derived envelope signal; (4) a nonlinear mapping operation which reduces the relatively large range of band levels (50-60 dB), to a narrower range consistent with the dynamic range of an electrically-stimulated implanted electrode (5-25 dB); and (5) a modulation operation which imposes the mapped envelope variations upon the pulsatile stimulation waveform amplitude.

As shown in Figure 7, each channel's operations can be viewed as an input process (with an equivalent gain, G_{in}) whose output is the input signal to the nonlinear mapping operation. The output of the mapping is adjusted by another gain (G_{out}) and converted to a stimulation current. Because of the nonlinear mapping, various pairs of G_{in}, G_{out} will produce a comfortable loudness sensation, but, for each of these pairs, the distribution of the band levels will be positioned differently with respect to the mapping function. One important aspect of fitting a CIS processor to a particular subject is the adjustment of the G_{in}, G_{out} pair.

Fitting procedures that specify the CIS processor parameters usually include psychophysical measures of the threshold level (THR) and the most comfortable loudness level (MCL) for each of a subject's implanted electrodes. In some cases, the manner in which the perception of loudness grows with increasing current level and statistical measures of the input speech signal are also employed. These measures are used to establish values for G_{in} and G_{out} that determine a suitable operating point.

The next three sections present a comparison of the CIS implementation and fitting procedures used by three research sites that are studying subjects' use of the Geneva Wearable Processor (GWP) implementing CIS⁴. We refer to the three implementations as the MIT/GWP, the Geneva/GWP, and the RTI/GWP processing schemes. Table I summarizes the operations of these three systems and Figures 8, 9, and 10 show the magnitude responses of their pre-emphasis, bandpass and lowpass filters respectively.

Table I. CIS Implementation Comparison

Operation	Geneva	MIT	RTI
AGC	peak limiting only	none	none
Equalization	HPF 1200 Hz (6dB/oct)	none	HPF 1200 Hz (6dB/oct)
Bandpass Filters	6 FIR, Hamming wind, log-spaced 150-6440 Hz - 6dB cross, -55 dB stop	6 FIR equiripple, log-spaced, 250-7000 Hz -6dB cross, -60dB stop	6 IIR, Hamming (6th), log-spaced, 300-5500Hz - 3dB cross
Envelope Measure	Quadrature (complex)	Quadrature (complex)	Magnitude (FWR)
Lowpass Filters	FIR, Hamming wind, -3dB@400, -55dB@1k	FIR, equiripple, -3dB@400, -60dB@1k	IIR, 4th order Butterwth - 3dB@400Hz
Map Input Level (Gin)	All channels @ 1% sat.	strongest channel@1%	set for comfort. loudness
Map Input Range	50-55dB	60dB	54dB
Mapping Function	Out= A*(In**0.001)+B	Out= A*(In**0.001)+B	Out= A*(In**-0.001)+B
Psychophysical Measures, Adjustments	Uses 500ms CIS signal, THR, Loudness growth, use 50% loudness point	Uses 50ms CIS signal, THR, MCL, Imax= k[MCL-THR] + THR	Uses 50ms CIS signal, THR, MCL. kTHR, kMCL
Stimulation Waveform	Cathodic 1st, 1667 Hz, 50 microsec./phase	Cathodic 1st, 2000 Hz, 31.25 microsec./phase	Anodic 1st, 2525 Hz, 32 microsec./phase
User Controls	none	Input Sens.	Input Sens, Output Vol.

4.1.1 The MIT/GWP Implementation

Referring to Figure 7. and Table I, there is no front-end processing in the MIT/GWP implementation. The microphone signal is applied to the band-pass filters after some amplification that is not frequency dependent. There is no automatic gain control applied before the filtering. Each of the filters is an equiripple, finite impulse response (FIR) design which spans the frequency range from 250 to 7000 Hz with log-spaced crossover frequencies at -6 dB, and -60 dB rejection in the stop bands (see Figure 9.). The envelope is derived using quadrature channels to form a complex signal whose magnitude is the measured envelope. This envelope is smoothed and limited in its frequency range by an equiripple low-pass FIR filter whose -3 dB bandwidth is at 400 Hz and is -60 dB at 1000 Hz (see Figure 10.). The output of this filter is applied to the mapping function at a level described in the fitting procedure below. The input range of the mapping is 60 dB, and the function is of the form: $Out = A*(In^{**}0.001) + B$ where the constants are used to fit the Out signal to the range from the THR current level to a maximum current (Imax) determined from the psychophysically measured MCL. This mapped quantity, (Out) is used to modulate a cathodic-first, bipolar pulse train with repetition rate of 2000 Hz and phase duration of 31.25 microseconds. Six of these envelope modulated pulsatile stimulation currents are applied to electrodes starting at the most basal, in the order 6-3-5-2-4-1. We have not observed a strong effect when using other stimulation orders (e.g. 6-5-4-3-2-1, or 1-2-3-4-5-6).

As mentioned earlier, the fitting procedure serves to define the G_{in} and G_{out} parameters based upon: (1) psychophysical measures of THR and MCL made for each subject's electrodes and (2) measures of level histograms made for each CIS processing channel. We use a subset of the TIMIT database⁵, sentences spoken by male and female talkers from eight dialect regions in the US, as the speech input to compute the band-level histograms. The sentence subset used consists of one male and one female talker from each of the eight dialect regions, each speaking ten sentences that are played at a comfortable listening level. G_{in} is set to be uniform across channels and is adjusted so that the highest 1% of the envelope levels generated by the TIMIT subset reach the clipping or saturating level at the mapping input for the highest energy channel (in our processor, channel #2)¹. When speech is applied to the processor at this level and the output of the mapping is allowed to range from a highest stimulation current of MCL to a lowest value of THR (envelope levels below -60 dB re: max input are set to zero), the perceived loudness of the CIS processor is above the level of comfort. We maintain the minimum output level at THR, and reduce I_{max} from MCL to a level determined by $I_{max} = k[MCL - THR] + THR$, where k is adjusted to produce a comfortable listening level. For the twenty Ineraid subjects we have fitted, the value of k ranges from 0.4 to 0.7. The computed mapping function uses an input dynamic range of 60 dB and an output range (for each electrode) of $20 * (\log_{10}(I_{max}/THR))$. This fitting procedure was used with the twenty Ineraid implanted subjects who are currently using the MIT/GWP processor.

4.1.2 The Geneva/GWP Implementation⁶

Again referring to Figure 7. and Table I, the Geneva/GWP system employs both an AGC operation and frequency shaping of the microphone signal before the band-pass filters. The AGC operation is set so that conversational-level signals are not affected by the AGC amplifier, but extremely loud signals are limited in energy so as not to cause uncontrollably loud stimulation. The output of the AGC amplifier is pre-emphasized by a high-pass filter whose -3 dB frequency is set to 1200 Hz and rolls off below this frequency at 6 dB/octave. This preprocessed signal is then filtered by six FIR band-pass filters spanning the range from 150 to 6400 Hz with log-spaced crossover frequencies at -6 dB. The stop band attenuation for each filter is -55 dB. The envelope of each filter output is measured using the same quadrature operation as the MIT/GWP system. Each envelope signal is then low-pass filtered with an FIR filter whose -3 dB point is at 400 Hz and is down by -55 dB at 1000 Hz. Each of the six smoothed envelope signals is multiplied by a separate gain value before the mapping operation, so that each of the envelopes reaches the peak input (saturation) level 1% of the time for conversational level speech input. Because the G_{in} varies across channels, the input envelope range that is mapped varies for each channel from approximately 50 to 55 dB. We will say more about this below. The mapping function is implemented with a table of 512 values of the function: $Out = A * (In * 0.001) + B$. The output range of this mapping table varies over a range described in the next section. The mapped output range is applied as amplitude modulation to a cathodic-first, bipolar pulse train with repetition rate of 1667 Hz and phase width of 50 microseconds. The update order of electrode stimulation is (starting at the basal end of the cochlea) 6-3-5-2-4-1.

The Geneva fitting procedure begins by adjusting the G_{in} for each channel based on: (1) psychophysical measures of THR and MCL and (2) level statistics measured using four-talker speech babble. The speech babble is constructed by starting with four-talker babble and then creating three new versions by delaying the original by three seconds three times. These four babble tracks are then mixed to form a composite babble of "sixteen" talkers. This composite recording is played in a small anechoic chamber at 75 dB SPL(A) through the earhook microphone of the processor used for the fitting. Using all of the processing described in the previous paragraph, histograms of the six envelopes, before the mapping operation, are measured. For *each* channel, the input mapping gain is set so that 1% of the input envelope values are at or above the saturation level. To determine the lowest point of the mapping input, the babble is turned off and histograms are made of the noise envelopes for each channel. For *each* channel, the band level that is higher than 99% of the noise levels is used as the lower boundary of the mapping function's input range and is mapped to THR. For this procedure, the mapping input ranges from approximately 55 dB for channel #1 to approximately 50 dB for the remaining channels. The adjustment of the G_{in} for each channel imposes a frequency weighting (at least on a channel by channel basis) that is much like pre-emphasis.

When the input is adjusted as described, the output range of the mapping must be converted to a range of stimulation currents that provide comfortable loudness for conversational speech. The peak stimulation current for each channel is chosen from psychophysical measurements of each electrode's loudness growth function using 500 millisecond bursts of the CIS pulsatile waveforms. The current which produces a 50% loudness level (from THR to MCL) is used as the upper limit of each channel's stimulation current. The minimum output current is the measured THR. If this produces a loudness which is too loud or soft, then the 40 or 60% loudness levels will be used to determine peak stimulation currents. In principal, neither output loudness nor input sensitivity need further adjustment by the subject while the processor is in use (especially given that the input level is limited by the AGC circuit).

4.1.3 The RTI/GWP Implementation

The RTI/GWP uses no input AGC but does employ the same pre-emphasis processing as described for the Geneva system: a high-pass filter with cutoff frequency at 1200 Hz and a roll-off of 6 dB/octave below that frequency (see Figure 7. and Table I). One of the standard band-pass filter configurations used by RTI is realized by six, 6th order Butterworth filters which are log-spaced over the frequency range from 300 - 5500 Hz, with crossover frequencies at the -3 dB points. The envelope of each band-pass filter output is measured by computing the magnitude of the waveform samples (referred to as full-wave rectification of the band-pass waveform) and smoothing this signal with a low-pass filter. The low-pass filter is implemented with a 4th order Butterworth filter with a -3 dB point at 400 Hz. The smoothed envelope is the mapping function input. The mapping function uses an input range of 54 dB and implements the function: $Out = A * (In^{**} - 0.001) + B$, with a 512 point table. The output value is used to modulate an

anodic-first pulsatile waveform with repetition rate of 2525 Hz and phase duration of 32 microsecond. The update order of electrode stimulation varies from subject to subject -- base to apex, apex to base, or alternating 6-3-5-2-4-1. The output range of the stimulation currents are based upon psychophysical measurements made during the fitting procedure.

The fitting procedure for RTI subjects is based upon psychophysical measures of MCL, and THR using a 50 millisecond burst of a CIS waveform (2525 Hz, 32 microsecond/phase). These two levels are used to compute the "A" and "B" of the mapping function so that the output currents range from MCL to THR. In contrast to the MIT and Geneva procedures, there is no attempt to specify the input range of envelopes that are mapped. Instead, with the output ranges as described, the input gain from the microphone is adjusted so that conversational speech provides stimulation currents that evoke comfortable loudness levels. There may be some slight adjustment of the output stimulation, after interaction in the laboratory, such that MCL and THR currents are reduced by an overall factor, K. That is, MCL is reduced to $K \cdot \text{MCL}$ and THR is reduced to $K \cdot \text{THR}$, where K ranges from between 0.8 to 1.0. Using this range of stimulation currents suggests that an RTI processor implementation probably operates at lower levels into the mapping than either the MIT or Geneva implementations.

4.2 Mapping Range and Channel Gain Studies

4.2.1 Range of Envelope Levels Mapped vs Speech Reception

The work described in this section explored the effect of reducing the range of input envelope levels mapped across channels ⁷. As discussed earlier, the CIS algorithm used by the MIT/GWP subjects maps a 60 dB range of envelope levels into a range of stimulation currents that varies depending on the DR of the particular electrode. In the work reported in this section, two modifications of the standard CIS processing algorithm were made. First, instead of using a constant G_{in} across channels that is based on the 1% clipping criterion in Channel 2 as described above, the G_{in} of each channel was adjusted separately so that all channels clipped 1% of their levels when the TIMIT sentences were played at a conversational level (similar to the G_{in} used by the UoG group).

With the CIS processor set in the above fashion, the mapping function of each channel was manipulated to reduce the range of envelope levels mapped as shown in Figure 11. In one set of experiments the range of envelope levels mapped was reduced by moving the low-level boundary of the mapping function's input range to the right (left column of Figure 11.). In a second set of experiments, the range of mapped levels was reduced by moving the high-level boundary of the range to the left (right column of Figure 11.). The position of the high or low-level boundary was selected to include a specific percentage of envelope levels rather than to include a specific input mapping range as shown in Figure 11. This can be seen in Table II where the low-level and high-level boundaries are specified as the percentage of levels lower than the boundary level. Thus, for a low-level boundary position of 80%, the mapped range in each channel begins at a level that is higher than 79% of all the levels in each channel. Similarly, a high-level boundary

position of 87% means that clipping begins at a level that is higher than 87% of all the levels in that channel.

Table II. Results of Limiting the High and Low Ends of the Mapping

Low-level Boundary Variation			High-level Boundary Variation		
Boundary Position		24 Consonant Score	Boundary Position		24 Consonant Score
Low	High		Low	High	
1	99	99%(1.86)	1	99	99%(1.86)
40	99	98%(3.73)	1	95	98%(3.73)
80	99	99%(1.86)	1	90	94%(3.73)
90	99	92%(1.86)	1	80	87%(4.56)
97	99	87%(5.43)	1	70	77%(2.28)
98.5	99	75%(6.59)	1	60	58%(6.59)
98.7	99	69%(2.28)			

Table II presents the effects of these two types of mapping manipulations upon subject I04's ability to identify 24 medial consonants spoken by a male talker/cite{iowa}. In the case where the high-level boundary level was reduced (the right half of the table), increasing the percentage of levels clipped, significant reductions in consonant scores began at the 1-80% condition. In a conversational setting, the subject could not detect a qualitative difference between the 1-99% and the 1-95% conditions, but did begin to hear increased levels of distortion at the 1-90% condition.

The left half of Table II shows the consonant identification scores for manipulations of the low-level boundary of the mapping function. In this case, the lowest 96% of the envelope levels must be discarded to reach a consonant score comparable to the condition of clipping the highest 20% of the envelope levels. Qualitatively, the subject preferred the 40-99% condition to the 1-99% condition in a conversational setting because of a reduction in the ambient background noise; voice quality seemed to be identical. As the low-level boundary is moved higher and a larger percentage of the envelope levels are discarded, the voice quality begins to suffer from "dropout effects." The subject began to detect this type of distortion when more than 40% of the envelope values were discarded, even though the consonant scores remain high.

To some extent, the effects of these manipulations are minimized by the controlled levels of the speech segments used in the consonant test. Nevertheless, it is clear that level variations at the high end of the range of input envelope levels are more important to represent than those at the low end of the range. Thus, given a limited input range to be mapped, one should adjust G_{in} to avoid clipping much beyond a 1% level.

4.2.2 Restricted Envelope Ranges Compared to Our Standard Mapping Ranges

Our standard CIS processor maps a 60 dB range of input envelope levels into an output current range that is specified for each channel. The subjective judgment of I04

(reinforced by the 24-consonant test results) that the lower 40% of envelope levels can be discarded with no reduction in sound quality, suggests that it may be possible to reduce the mapped range of envelope levels without compromising performance. If the length of the table representing the mapping function remains the same, a reduction of this range would mean an increase in the amplitude resolution. Alternatively, maintaining the same amplitude resolution and reducing the range of mapped envelope levels would mean the mapping function could be represented in a shorter table thereby reducing the memory required by the CIS program.

The range of envelope levels that correspond to 40-99% boundary positions shown in Table II can be determined from the cumulative distribution of envelope levels computed for each channel before the mapping stage and shown as tables in Figure 12. For example, the cumulative level distribution for channel #3 shows that the 99% high-level boundary corresponds to -10 dB and the 40% low-level boundary corresponds to -46 dB. This means that a range of 36 dB is required to map the envelope levels within the 40-99% boundary positions in channel #3. The same process applied to the other channels gives the following ranges for channels #1 through #6: 23, 34, 36, 30, 29 and 32 dB.

In our standard CIS system, G_{in} is constant across channels and set to clip 1% of the envelope levels in channel #2. For the case shown in Figure 12, G_{in} (see Figure 7.) would be set to 5 dB because the cumulative level distribution for the TIMIT sentences show that 99% of the levels in channel #2 are 5dB below the channel's maximum (16 bit) level. This means that the input ranges for channels #1 through #6 that would position the lower-level boundary at 40% of the channel's cumulative level distribution are: 29, 34, 41, 39, 40 and 44 dB re the channel's maximum level.

If measurements in additional subjects show similar results to those of I04, they would suggest reducing our standard 60 dB mapping range to one of 40-45 dB or less without adversely affecting speech reception and sound quality. However, three factors work to minimize the actual reduction realizable without a decrease in sound quality: (1) the CIS implementation used with the GWP system does not include an AGC to reduce the variations in overall sound levels encountered in the real world, (2) the medial consonant test used to measure speech reception for the various level boundary conditions does not include a wide range of overall amplitude within or across the various speech tokens and (3) the TIMIT sentences used to produce the cumulative level distributions are also well controlled in terms of overall level.

4.2.3 Choices for Channel Gain Equalization

We have mentioned two strategies for setting G_{in} using a standard set of speech materials (the TIMIT sentences in our case). In one, G_{in} is the same across all channels and is set such that 1% of the levels will be clipped in the channel with the greatest energy. Another strategy is to set G_{in} separately for each channel. In this case the G_{in} of each channel is set such that 1% of the levels are clipped in each channel. The first method maintains the spectral shape of the speech because the gain across channels is the same. In the second

method, the gains vary across channels and the spectral shape is distorted. When the TIMIT sentences are used to set these individual gains, a high frequency emphasis is produced because the G_{in} needed at each channel to move the 99% point of the cumulative level distribution is, in general, greater for the high-frequency channels (e.g., 11, 5, 10, 14, 16, and 17 dB for channels #1 - #6 respectively as can be seen from Figure 12.).

To determine whether the equalization on a channel basis enhances speech reception even though the overall speech spectrum suffers a constant distortion, we tested four subjects for medial consonant reception when using our six-channel processor. In one case each channel's gain was adjusted so that all six mapping inputs saturated 1% of the time when the TIMIT speech materials were used as input. In the second case an equal gain was used for all six channels such that the strongest channel's output (channel #2) saturated 1% of the time. For our CIS processor implementation, which uses a 60 dB mapping band level, the results for four subjects are shown in Table III. In each of the four subjects, average percent-correct scores for the 16 consonants spoken by a male talker were higher for the case of equal gain. However the limited amount of data shows significance for only three of the four subjects as indicated.

Table III. Scores for Two Choices of Input Gain (G_{in})

Subject	All at 1%	Only channel#2	comments
S02	50%(12.4)	56%(8.5)	significant at 0.1
S11	45.2%(4.5)	51.6%(4.6)	significant at 0.25
S01	73.8%(3.8)	79.5%(8.4)	significant at 0.05
S27	29.4%(14.2)	37.6%(13.9)	not significant at 0.1

Although the evidence is not strong, it certainly does not indicate a clear advantage for the equalized gain condition. Intuition would suggest that if the input band levels will accommodate the range of envelope values required for good speech reception (as discussed earlier), then there is no advantage to the introduction of a spectrum-distorting equalization. However, if the band levels are highly reduced, then the distortion may buy some reception advantage.

4.3 Loudness Growth and CIS Mapping Studies

The initial goal of the work we report in this section was to design and test CIS mapping functions that restore normal loudness growth (NLG) of tones for implantees. This work⁸ also led us to revisit methods used to define the maximum stimulation currents (I_{max}) produced by mapping functions used in CIS processors. One method tested defined I_{max} for an electrode using a criterion based on a fraction (α) of the electrode's dynamic range. Mapping functions using this method are called " α functions". Another method tested used a loudness-based criterion for selecting I_{max} . Mapping functions based on this method are referred to as "L functions".

Having investigated the procedure for defining the peak stimulating currents produced by CIS mapping functions, we developed a method that defines the shape of mapping functions themselves so as to restore the normal growth of loudness in cochlear implant

users for tones. This method required measuring: (1) LG functions for each of a subject's intracochlear electrodes using electrical stimulation and (2) the LG function of normal-hearing subjects using acoustic tones. Using both growth relationships we defined the transformation required between the acoustic signal and stimulation current to provide normal growth of loudness for cochlear implant users. This transformation determines a new set of mapping functions that we designate by the term "NLG mapping functions".

In the following sections, we: (1) present LG measurements conducted with seven Ineraid cochlear implant subjects, (2) describe the two methods used to define the maximum stimulation currents used for logarithmic mapping functions, (3) report speech reception performance obtained for these two logarithmic mapping functions, and (4) report speech reception performance for sound processors using the NLG mapping functions.

4.3.1 Methods

Subjects

Seven Ineraid subjects (3 female and 4 male) participated in this study. All of them are postlingually deafened and had used the Ineraid processor for more than five years. At the time of this study, subject S22 had been using the Innsbruck Research wearable processor⁹, for 17 months. For at least one year, the other subjects had been using the Geneva Wearable Processor with a version of the CIS sound processing strategy developed at MIT¹⁰.

Loudness Growth Measurements

LG measurements were conducted for each of the seven subject's electrodes using an absolute magnitude estimate (AME) procedure, that requires the subject to assign a number to the loudness perceived for a given stimulus. Before conducting the AME with a cochlear implant subject, we measured the threshold (THR) and the maximum comfortable listening level (MCL) for each electrode to determine its dynamic range for electric stimulation. All the measurements conducted in this study were obtained using 300 ms, biphasic pulse trains, with parameters like those of the pulse trains used as carriers by the CIS strategy employed in this study (2000 pps per electrode, 31.25 μ s/ph, except for subject S05, where 40 μ s per phase was used).

The THR and MCL measurements were conducted using one 300 ms train for each stimulus amplitude. THR was measured using a 3-Alternative-Forced-Choice, "one up/two down" adaptive procedure that converges on a stimulus level where approximately 70% of presentations are detected¹¹. MCL was defined as "the loudest sound you can listen to comfortably for a long period of time" and was measured by the method of limits.

Absolute magnitude estimates were obtained for 20 stimulus amplitudes distributed linearly over the range from THR to MCL. The 20 amplitudes were presented eight

times, all in pseudo random order. The “randomization” of amplitudes was accomplished by randomly drawing (without replacement) 8 lists from a pool of twenty lists. The twenty lists were selected from a larger number of lists (each containing the twenty stimulus amplitudes in different order) to comply with the following constraints: no amplitude immediately precedes any other amplitude more than twice and no two successive amplitudes differ by more than 60% of the dynamic range.

The subject was instructed as follows: “We will randomly present 20 sounds of different amplitudes. When a sound is presented, a box drawn on the screen will light. For each amplitude, we will present two sound bursts. You should describe how loud they are by assigning a number to them. You may use any positive number (e.g. 3000, 500, 70, 0.6, 0.04, etc.). Answer ‘0’ if you do not hear a sound. Do not worry about consistency. Simply try to match an appropriate number to each tone regardless of what you may have called the previous stimulus.”

We determine the LG function $L_e(I_e)$, for each electrode, from the 8 magnitude estimates obtained for each of the 20 different stimulus amplitudes using a least-square fourth-order polynomial fitting routine (Matlab™ Toolbox):

$$L_e(I_e) = a_{0,e} + a_{1,e} \cdot I_e + a_{2,e} \cdot I_e^2 + a_{3,e} \cdot I_e^3 + a_{4,e} I_e^4 \quad \text{Eq. 1}$$

where L_e is the loudness produced by stimulating electrode “e” at stimulus amplitude I_e and $a_{i,e}$ are the coefficients of the polynomial fit.

Sound Processors

We used the laboratory digital signal processing facility as described in a previous report¹⁰ to conduct the LG measurements and to test different mapping functions with the CIS strategy in the laboratory. Field tests were conducted using the GWP sound processor. Details of the CIS strategy used in this study are as described in an earlier section.

α and L Mapping Functions

At the input of the mapping function, a gain (G_{in}), determines the range of envelope levels that will fall within the 60 dB input range. G_{in} is set so that 1% of the envelope levels will be clipped in the channel with the most energy (channel 2) when the TIMIT (Fisher et al. 1986) data base of speech materials is played at a conversational level. The standard equation of the mapping functions implemented for this strategy is

$$I = a \cdot x^p + b$$

where x is the envelope amplitude within a frequency band; I is the amplitude of the electric stimulation for the associated electrode;

$$a = \left(\frac{I_{\max} - I_{\min}}{x_{\max}^p - x_{\min}^p} \right);$$

$$b = I_{\max} - a \cdot x_{\max}^p;$$

with $[x_{\min} : x_{\max}]$ and $[I_{\min} : I_{\max}]$ respectively, the input and output ranges of the mapping

function. The parameter p defines the shape of the mapping function. When $p = 0.001$ the mapping approximates the logarithmic function used in many of today's sound processing systems.

The output range of the mapping function can be defined by setting I_{min} near THR and selecting I_{max} to produce a most comfortable listening level. The initial method used to define each channel's I_{max} was to use a proportion (α) of the electric range defined by the channel's THR and MCL ($I_{max} = \alpha * (MCL - THR) + THR$). For a given mapping function shape (e.g., logarithmic) α can be adjusted for a most comfortable listening level. The I_{max} currents defined in this way characterize the logarithmic mapping functions that we designated by " α functions".

Figure 13. shows LGFs for electrodes 3 and 6 in subject S04. If, for instance, the I_{max} currents are defined to match 80% of the dynamic range, we can see from the LGFs that the single channel loudness values associated with these I_{max} currents are different across electrodes (e.g., 66 and 47 for electrodes 3 and 6 respectively). These differences are due to differences in the shape of the LG function and the fact that the MCL at electrode three produces a loudness sensation that is different than that produced by an MCL stimulus level at electrode six.

Another method for selecting I_{max} , for each mapping function is to use a constant-loudness criterion, L_2 , that should be produced by the I_{max} of each channel. This loudness criterion is adjusted to provide a most comfortable listening level. The I_{max} s obtained by this method are used to define the logarithmic mapping functions called "L functions".

NLG Mapping Functions

The goal of restoring NLG to a cochlear implant subject requires a mapping function that, for a specified acoustic input range, delivers stimulus levels producing the same growth in loudness sensation for an implant user as that experienced by normal-hearing listeners. To compute NLG mapping functions, we compare the electrical LG of each electrode $L_e(I_e)$, to the normal hearing LG obtained for acoustic tones, in order to define the relationship between the two modalities. The normal LG for pure tones has been described by the power law:

$$L = k * P^\alpha \quad \text{Eq. 2}$$

where P is the acoustic pressure of a tone. The exponent α has been defined by Stevens (1955, 1957a) as equal to 0.6 based on magnitude estimation data he obtained when subjects were instructed to match the ratios between the numbers they assigned to pairs of stimuli to the ratios between the loudness of the sensation elicited by those stimuli. An exponent of 0.6 implies that a 10dB increase in tone level will double the perceived loudness.

Hellman and Meiselman (1988) also recorded magnitude estimation for loudness, but simply instructed subjects to assign numbers based on loudness (they did not

introduce the concept of ratios). Their data showed a mean exponent of 0.46. Because the exponent of the power function depends on the method used to measure loudness, we decided to measure α for acoustic tones in five normal hearing subjects using the same methods employed to measure LG functions for electric stimulation in implant subjects. The details of this work can be found in a previous report (Eddington et al. 1997). In this study we obtained an exponent α with a mean of 0.42, a minimum of 0.32 and a maximum of 0.55. Because of the across-subject variability, we defined three sets of NLG functions corresponding to α 's of 0.32, 0.42, 0.55, implemented them on the Geneva Wearable Processor, and evaluated them with subjects.

To compute mapping functions restoring NLG, we defined a relation between acoustic level and electric amplitude based upon data obtained acoustically from normal hearing subjects and electrically from cochlear implant subjects. From the power law describing loudness growth of normal hearing subjects, we can express the logarithm of the ratios between the loudness magnitude estimates and the reference loudness:

$$\log\left(\frac{L_e}{L_2}\right) = \alpha * \log\left(\frac{P_n}{P_2}\right) \quad \text{Eq. 3}$$

where L_e is the LG function for electrode e that is associated with channel n and P_n is the sound pressure of a tone at the center frequency of channel n .

Because the sound pressure (P_n) for a tone at the center frequency of a channel is linearly related to the envelope amplitude (E_n) for the same channel, we can write:

$$\log\left(\frac{L_e}{L_2}\right) = \alpha * \log\left(\frac{E_n}{E_2}\right) \quad \text{Eq. 4}$$

Eq. 4 expresses how the loudness produced by electrical stimulation should vary as a function of the band envelope level. Because Eq. 1 relates L_e to electric stimulus level, we have all the information needed to define $I_e(E_n)$, the NLG mapping functions.

Speech Reception Tests

We obtained speech reception measures without lipreading and in quiet for CIS processors using the three mapping functions described above. The subjects' ability to identify monosyllabic words was measured using the NU6 test and their ability to identify consonants in an /aCa/ context was measured using the Iowa 16 and 24 consonant tests ¹².

4.3.2 Results

Speech reception with logarithmic mappings

Subjects had been using a CIS strategy implemented with the α logarithmic mapping functions in daily life with the Geneva Wearable Processor (GWP) for about a year. Then they tested a CIS strategy implemented with the L logarithmic mapping functions (where the I_{max} s are based on an equal loudness criterion) for a few weeks. The left panel in

Figure 14. shows percent correct consonant identification for the two implementations using logarithmic mapping. The bars in white represent the scores obtained with the α logarithmic mapping functions and the bars in gray represent the scores obtained with the L logarithmic mapping functions. Each bar represents the mean score computed from at least 10 randomized presentations of the 24 consonants (or 16 consonants for subject S02, S05, S15). The individual data show that all subjects score higher using the L logarithmic mapping functions.

The right panel in Figure 14. presents the identification scores for monosyllabic words. Each bar represents the mean identification score computed from the presentation of 1 or 2 lists of 50 monosyllabic words. These scores also show a consistent advantage for the L logarithmic mapping functions. These consonant and single-syllable word scores are consistent with the clear preference articulated by all subjects for the L logarithmic mapping function processors.

Loudness growth

Figure 15. presents the loudness growth obtained by subject S01 for his most apical electrode. The stars represent the medians of the 8 loudness estimates obtained for each of the twenty amplitudes presented. The error bars represent the interquartiles computed on these same loudness estimates. On this graph the fitted loudness growth function ($L_e(I_e) = a_{0,e} + a_{1,e} \cdot I_e + a_{2,e} \cdot I_e^2 + a_{3,e} \cdot I_e^3 + a_{4,e} \cdot I_e^4$) is also drawn, with the 95% confidence limits for the prediction shown as two dotted lines. The multiple correlation coefficient R^2 is also given.

Figure 16. shows the LG functions, $L_e(I_e)$, for each electrode of the 7 subjects studied. The maximum loudness estimate assigned to the MCL is different across electrodes. Note also that the shapes of the loudness growth functions are different across subjects and across electrodes. For a logarithmic mapping function to produce NLG on a single electrode, the loudness growth functions measured on that electrode would need to be exponential. Figure 17. compares the loudness growth functions measured on all electrodes for two subjects (S01 and S15; solid lines) to such an exponential growth function. Note that the growth functions for subject S01 are clearly different from the exponential that has been used by others to describe the growth of loudness in electric stimulation¹³. Some of S15's growth functions however, are much closer to the exponential.

NLG Mapping functions

We computed sets of NLG mapping functions from the loudness growth functions $L_e(I_e)$ measured for each electrode using exponents of 0.32, 0.42 and 0.55 for most subjects. Figure 18. presents the mapping functions for each electrode for the "preferred" exponent and the logarithmic function (diagonal line) for comparison. Depending on the electrode and the subject, NLG mapping functions can be very different from the logarithmic mapping functions. Table IV shows the exponent preferred by each subject (when more than one was tested) and, in some cases, comments relating to their preference.

Table IV. Power Law Exponents Preferred by Seven Subjects

Subject	α Power Law Exponent Preferred	Comments
S01	0.32	0.42 and 0.55 soft sounds too soft
S02	0.32	0.42 and 0.55 soft sounds too soft
S04	0.32	
S05	0.42	doesn't like 0.55; 0.32 not tested
S15	0.32	0.32 lower scores, 0.55 too soft
S20	0.42	only 0.42 tested
S22	0.42	0.32 more background noise; 0.55 closer to the Ineraid

Speech reception with NLG mappings

While the new mapping functions are designed to restore the normal growth of loudness for single tones, we also evaluated the effect of these mapping functions on speech reception. three weeks. Figure 19. shows the mean identification scores computed for at least 10 randomized presentations of the 24 (or 16 for subject S05) consonants for five of the seven subjects for whom NLG mapping functions had been determined and had used the new mapping for at least three weeks before testing. The bars in gray represent scores obtained with the L logarithmic mapping functions and the bars in black represent scores obtained with the NLG mapping functions. For consonants, individual scores tend to be higher for the NLG mapping functions, but the group difference is not statistically significant.

The right panel of Figure 19. shows scores for single-syllable word recognition. Notice that the two subjects (S01 and S02) who possess NLG mapping functions that are the most different from logarithmic mapping functions show scores that are higher for the NLG mapping processors. The differences in overall scores are not statistically significant. Subjects S01 and S20 preferred using the NLG mapping functions in all listening conditions. The others used the logarithmic mapping function in quiet environments and the NLG mapping function in noisy environments. The comparison of monosyllabic identification scores obtained in quiet with the three different mapping functions shows that the effect on speech reception from changing the shape of the map (NLG vs. L logarithmic), is smaller than that obtained from optimizing the selection of I_{max} (α logarithmic vs L logarithmic).

4.4 Longitudinal Speech Reception Comparisons of Ineraid and CIS Processors

A number of investigators have studied the speech reception of Ineraid implantees as they moved from their commercial Compressed Analog (CA) processor to a version of the Continuous Interleaved Sampling (CIS) processor in the laboratory and in the field ¹⁴⁻¹⁷. Wilson has noted that virtually every Ineraid subject fit using their laboratory CIS processor demonstrated significant improvement in speech reception during the several-day fitting session ¹⁸.

The Geneva Group also reported that most Ineraid users experience a significant improvement in speech reception at the initial fitting of a CIS processor ¹⁹. Figure 20., for instance, shows CA and CIS scores for 22 subjects obtained with a 14-item consonant test. The CA processor was the commercial Ineraid system that each subject had been wearing for at least six months and the CA score for each subject represents the mean of the last five scores obtained with that system. The CIS processor score also represents the mean of five tests, each conducted on a different day. Except for the testing and the short conversations related to adjusting the CIS system, the subjects did not have an opportunity to practice and learn with this system. While at least five subjects do not show significant improvement with the CIS processor, the majority probably demonstrate functional improvements.

Our experience with Ineraid subjects moving from their CA system to a CIS sound processor is summarized in Figure 21. and does not show a majority of subjects demonstrating an immediate improvement. The top panel of Figure 21. plots the scores achieved on the Iowa 16-Consonant test ²⁰ by 18 subjects on the day they were fit with a wearable CIS sound processing system. The CA scores were obtained using the subjects' Ineraid processors (at least six months experience). The CIS processor was implemented using the Geneva Wearable Processor (GWP) running software developed by us (MIT/GWP) ¹. Note that unlike the experience of Wilson and the Geneva Group, at least half of our subjects do not demonstrate an immediate improvement.

The bottom panel of Figure 21. compares the same CA scores with CIS scores obtained after each of the subjects had worn the CIS system for at least six months. Even with this added experience with the CIS system, at least seven (or 39%) of the subjects do not demonstrate marked improvements in speech reception (although virtually all express a strong preference for the CIS system and none will return to their Ineraid system).

A number of possible reasons for this apparent difference between subject performance at MIT and that reported by Geneva and RTI have been suggested (e.g., implementation, fitting methods, testing methods, and small numbers of subjects). One intriguing possibility is that because the developer of the Ineraid system fit the subjects in Boston, they received better optimized systems than other Ineraid subjects. Figure 22. presents mean data for Ineraid and CIS systems from Geneva (14-consonant test) and MIT (16-consonant test) that relates to this possibility. Notice first the difference between the Geneva (40.8%; $\text{stderr: } 4.1$) and MIT (51.2%; $\text{stderr: } 3.6$) mean scores for the Ineraid (CA). This difference is consistent with the hypothesis that one reason the

MIT subjects do not show the improvements in speech reception reported by RTI and Geneva is that the scores used as a basis for judging improvement are, on average, higher. Note also that the mean performance (at the time of fitting) of the MIT group of subjects using the MIT/GWP CIS processor is at least as high as the Geneva subjects. Because mean scores of information transfer computed from consonant recognition results collected for 11 subjects by RTI are similar to those collected for 15 subjects in Geneva²¹ and because the mean consonant scores collected with CIS processors at the time of fitting are similar to those reported by Geneva¹⁹, it is probable that a difference in performance does not exist between any of these three versions of CIS processing.

5 Bilateral Studies

The overall objective of our bilateral stimulation research was to determine if and how binaural advantages can be realized with a pair of cochlear implants. A fundamental step toward this goal is the measurement of binaural sensitivity. This section of the final report contains the results of binaural psychophysical experiments conducted with a single bilaterally implanted subject. One question we addressed is whether there is binaural sensitivity at a pitch-matched electrode pair. A second question is whether binaural sensitivity is worse for an electrode pair that is pitch mismatched.

Our primary assessment of binaural sensitivity comes from measurements of the subject's sensitivity to Interaural Time Differences (ITDs) in binaural psychophysical experiments. While we also measure Interaural Level Difference (ILD) sensitivity, the significant monaural sensitivity to level changes complicates the use of ILD sensitivity as a measure of binaural sensitivity. Because of the very poor monaural sensitivity to time delays, ITD sensitivity provides a straightforward assay of binaural sensitivity.

We found sensitivity to (perceptually-relevant) ITDs on a pitch-matched, interaural electrode pair. For a pair of electrodes which were pitch-mismatched, we were unable to observe any ITD sensitivity. This is consistent with data from normal hearing listeners that shows a limited interaural frequency offset for which ITD effects can be seen ²².

These results give evidence that it may be possible to exploit binaural sensitivity to provide binaural advantages to cochlear implant users. The results also indicate that the choice of electrode pair affects the binaural sensitivity and, therefore, this choice should be a consideration in attempts to provide binaural advantages.

5.1 Methods

5.1.1 Subject

The subject was a 72-year-old woman with an Ineraid implant (right cochlea) and a Clarion implant (left cochlea) who has been under the care of the Warren Otologic Group of Warren, Ohio. She apparently had normal hearing as a child, but at age 25 she noticed the onset of a bilateral hearing loss. Her hearing loss progressed bilaterally and she became deaf at age 44. At age 59, she received the Ineraid cochlear implant and used that system on a daily basis until age 70, when she received the Clarion implant in an effort to improve her hearing using new technology. Since then, she has been a full-time user of the Clarion implant alone. When she first received the Clarion implant, the subject attempted to use both Clarion and Ineraid implants simultaneously -- with

unmatched processors and electrode pairs -- for a brief period, but found the sensations confusing.

5.1.2 Stimulus Waveforms

Trains of fixed-amplitude, biphasic pulses were delivered to both the Clarion and the Ineraid electrodes. In the case of the Clarion implant, we report the requested current amplitude in quotes (i.e., "μ App") because the current actually delivered to the subject depends on a number of factors (including the electrode impedance) due to the nature of the implanted current source. A given "μ App" value will produce the same amplitude current (across a set load) as a given Clarion Clinical Unit (SCLIN 3.1 supplement) with half its value (e.g., 551.8 μ App" is equivalent to 275.8 Clinical Units). The amplitudes reported for the Ineraid implant are the levels actually delivered and, therefore, are not in quotes (i.e., μApp).

All electrodes were stimulated in a monopolar configuration. Stimulation of the 7th medial Clarion electrode and its far-field ground is indicated by "7MC." Stimulation of the 2nd Ineraid electrode with its far-field ground is indicated by "2I." For both devices, electrode 1 represents the most apical electrode. All stimuli were trains of 300ms duration, biphasic, cathodic-first pulses (76.9μsec phase duration). These waveform parameters differed across implant systems by less than 0.1%. During monolateral and non-simultaneous (alternating) bilateral stimulation, the pulse repetition rates were 100pps or 813pps. During simultaneous stimulation, the repetition rate was always 100pps.

ITDs were imparted by delaying the stimulation of the Ineraid electrode relative to the stimulation of the Clarion electrode (see Figure 23.). The total delay was a fixed hardware delay ($2.19 \mu\text{s} \pm 0.15 \mu\text{s}$) plus a programmable software delay, but in this report all ITDs will be given in terms of their software delay only.

The cochlear implant analog of Interaural Level Differences (ILDs) was produced by adjusting the Ineraid amplitude while holding the Clarion amplitude at a comfortable loudness level. The term ILD will refer to the Ineraid amplitude deviation from the Ineraid amplitude which produced a centered, fused (unitary) percept (with ITD equal to zero). In cases where a fused percept was not obtained, the ILD refers to a deviation from the Ineraid amplitude that produced a sensation equal in loudness to that of the Clarion comfortable stimulus level (with ITD equal to zero).

5.1.3 Preliminary Tests

The threshold and the uncomfortable loudness (UCL) levels were used to define the dynamic range for each test electrode. Absolute thresholds were measured using a three-interval, three-alternative, forced-choice task. A two-down, one-up, adaptive procedure was used to adjust the amplitude of the monolateral stimuli. Uncomfortable loudness

(UCL) was assessed with a one-interval, two-alternative, forced-choice task (one-down, one-up, adaptive) designed to find the monolateral stimulus amplitude that resulted in a sensation level on the verge of being uncomfortable.

A centering test was conducted with simultaneous, bilateral stimulation (100pps; ITD = 0 μ s) for electrode pairs in which such stimuli resulted in a fused percept. The amplitude of the Clarion stimulus was held constant at a comfortable level and the Ineraid stimulation amplitude was adjusted in order to produce a centered sensation. A one-interval, two-alternative, forced-choice (one-up, one-down, adaptive) procedure was employed. The result was used to determine the center of the range of Ineraid stimulus amplitudes to be used in the lateralization test. For an electrode pair in which the stimuli were not perceptually fused, the stimulus presentation method described above was used, but the subject was instructed to balance the simultaneous loudnesses of the two sound sensations.

Additional informal tasks were used to gather anecdotal information about fusion and lateralization. These tasks included (1) asking the subject questions about the sound sensations; (2) having the subject draw the perceived location of the sound sensations on a schematic picture of a head; and (3) the interactive use of a lateralization scale.

5.1.4 Pitch Ordering Test

This test was used to rank order the electrodes by the pitch sensations they produced. We assumed that electrodes eliciting similar pitch sensations (across ears) were located at similar cochleotopic positions. Based on the results of this test, cochleotopically matched (or mismatched) electrode pairs could be selected for further study in binaural psychophysical experiments. Before performing the interaural pitch ordering test, the subject was trained and tested with monolateral stimuli.

Each run of the test consisted of 25 presentations of the stimulus pairs. In the interaural test, two electrodes (at loudness-balanced amplitudes) were selected, one from each ear. In the monolateral test, two electrodes (at loudness-balanced amplitudes) were selected from the same ear. In the test, two boxes were drawn on the subject's display screen. Each presentation consisted of two stimuli presented in sequence. In the interaural test, the left box was lit during the presentation of the first (left ear) stimulus and the right box was lit during the presentation of the second (right ear) stimulus. In the monolateral test, all stimuli were presented to a single ear and the order of the two stimuli was randomized. For each, the subject's task was to type the number of the box (1 or 2) that was lit during the sound sensation that was higher in pitch. The test was a two-interval, two-alternative forced-choice task with stimuli that were fixed during each run. The stimulus waveforms were 813pps biphasic pulse trains presented non-simultaneously.

The results of the interaural pitch ordering test were expressed in terms of the percent of the presentations in which the subject judged the Clarion electrode as higher in pitch than the Ineraid electrode, $P(C>I)$. We classified the Clarion electrode as discriminably higher

in pitch than the Ineraid electrode (i.e., pitch-mismatched) if $P(C>I) \geq 76\%$ and as discriminably lower in pitch (i.e., pitch-mismatched) if $P(C>I) \leq 24\%$ for 25 presentations. For results between these bounds, the electrodes were classified as indiscriminable (i.e., pitch-matched). At the bounds, the sensitivity index, d' , equals 1.0 and for 25 presentations the probability in the upper or lower tail of the binomial distribution with $P=0.5$ is 0.0073. For 50 presentations, the upper and lower bounds of discriminability are 68% and 32%, respectively, and the probability in the upper or lower tail of the binomial distribution with $P=0.5$ (at the bounds) is 0.0077.

5.1.5 Lateralization Test

This test was designed to measure perceived locations (lateral positions) for a set of ILDs and ITDs on a given electrode pair. For each presentation, the subject assigned a number from a lateralization scale that represented the lateral position of that sound sensation. The scale range was from 1 to 7 where 1 represented a sound sensation at her left ear, 4 represented a centered sound sensation, and 7 represented a sound sensation at her right ear. The numbers in between allowed her to indicate intermediate positions. The subject was also instructed to indicate "None of the Above" (NA) if she could not assign a (single) number to the perceived location (e.g., if she heard more than one sound sensation during a single presentation). Thus, this was a one-interval, eight alternative (1, 2, 3, 4, 5, 6, 7, NA) test. A stimulus set was made up of 15 bilateral stimuli (5 ILDs and 3 ITDs). The ILDs were selected to elicit sensations at locations covering the range from the far left to the far right and were determined from the informal tests. The ITDs (0 μ sec, 300 μ sec, and 600 μ sec) were chosen to span the range of ITDs relevant in real world situations. On two days of testing with electrode pair 7MC/2I, a total of 33 of the randomized stimulus sets were presented.

5.1.6 Relative Loudness Test

This test was designed to determine if ITD had any effect when bilateral stimulation produced two separate (unfused) sensations and the lateralization test was, therefore, inappropriate. In this test, the subject was asked to assign a number to indicate the relative loudness of the two simultaneous sound sensations. The number 1 indicated that the sound sensation in the left ear was considerably louder than the sound sensation in the right ear. The number 4 indicated that the sound sensations were loudness balanced across the ears. The number 7 indicated that the sound sensation in the left ear was considerably softer than the sound sensation in the right ear. The numbers in between allowed her to indicate intermediate relative loudnesses. This one-interval, eight-alternative (1, 2, 3, 4, 5, 6, 7, NA) test used a stimulus set of 15 stimuli (5 ILDs and 3 ITDs). On one day of testing with electrode pair 3MC/2I, a total of 10 of the randomized stimulus sets were presented.

5.2 Results and Discussion

5.2.1 Preliminary Tests

Table V shows the threshold and UCL test results and the dynamic range calculations for electrodes 2I and 7MC with a repetition rate of 100pps. These tests were not performed on electrode 3MC.

Table V. Results from Preliminary Tests

Test	2I	7MC
Threshold	296.3 μ App	247.2 μ App"
Uncomfortable Loudness Level	826.3 μ App	650.4 μ App"
Dynamic Range	8.9dB	8.4 μ dB"

Table VI. Clarion Judged Higher in Pitch – Percent (# of presentations)

	Clarion el:	1MC	2MC	3MC	4MC	5MC	6MC	7MC	8MC
Ineraid el:									
1I					8(25)	68(50)	84(25)		
2I						20(25)	56(25)	48(25)	76(25)
3I					0(10)				
4I									
5I					0(10)				

The centering test with 7MC fixed at 551.8 μ App" gave a centered percept when the electrode 2I amplitude was 605.4 μ App. The simultaneous loudness balancing test with 3MC fixed at 408.8 μ App" gave a loudness balanced percept when the electrode 2I amplitude was 555.7 μ App.

5.2.2 Pitch Ordering Test

The results of the pitch ordering test (Table VI), are expressed in terms of the percent of the presentations in which the Clarion electrode was judged higher in pitch than the Ineraid electrode, P(C>I). When scanning from left to right across the table in a given row, this percentage generally increases. For example, the subject answered that 4MC was higher in pitch than 1I on only 8% of the presentations while she judged the more basal 6MC higher in pitch than 1I on 84% of the presentations. This trend is consistent with cochleotopic order.

Electrode pairs 7MC/2I and 6MC/2I, were classified as pitch indiscriminable. These pitch indiscriminable (i.e., pitch-matched) electrodes were taken as potentially good pairs for examining ITD sensitivity. Electrode pairs 5MC/1I and 8MC/2I were classified as bordering on discriminable. The other five electrode pairs tested were classified as discriminable (i.e., pitch-mismatched).

5.2.3 Electrode Pair 7MC/2I

Measurements of perceived lateral position as a function of ILD and ITD for electrode pair 7MC/2I are shown in Figure 24. Consider first the results with ITD = 0 μ sec. As might be expected, increases in the right electrode stimulus amplitude cause the mean lateral position of the sensation to move clearly toward the right. That is, lateral position is sensitive to ILD.

The results of Figure 24. also show lateral position sensitivity to ITD. For electrode 2I stimulated at 616 μ App, the mean lateral position of the elicited sensations shifts left as the right electrode stimulus is delayed. Similarly, the sensations produced by stimulating electrode 2I at 605 μ App (centered amplitude), tend to shift left as the right electrode stimulus is delayed. These shifts are significant ($\alpha < 0.01$ on a t-test) and demonstrate ITD sensitivity for these amplitudes. These results are consistent with normal hearing ITD effects on lateralization and are evidence of binaural sensitivity.

At some Ineraid amplitudes a significant ITD effect is not observed; these results can be understood in light of research with normal hearing listeners. For electrode 2I at 580 μ App, the ILD alone causes the percept to be lateralized to the left ear, leaving no room for an additional leftward shift with non-zero ITD. For electrode 2I stimulated at 627 μ App and 650 μ App, the lack of a significant ITD effect is consistent with results in normal hearing (for click stimuli) where large ILDs can greatly reduce the effect of ITDs on lateral position ^{23, 24}.

The subject answered "None of the Above" (NA) in 12.2% of her responses. While we did not routinely ask for detailed descriptions of the sensations eliciting NA responses, it is our impression that a significant number of the NA responses indicated the presence of two images.

5.2.4 Electrode Pair 3MC/2I

Electrode pair 3MC/2I was chosen as a pair that was cochleotopically mismatched based on the pitch ordering test. We hypothesized that if the ITD sensitivity observed with 7MC/2I were truly based on binaural mechanisms, this pair (3MC/2I) should not show a sensitivity to ITD. Qualitatively, the sensations produced by bilateral stimulation of this electrode pair were very different than those produced with 7MC/2I. Rather than a single, fused sensation, she heard two separate sound sensations with different pitches. She could not perform the lateralization test with this electrode pair, so we used the relative loudness procedure (see earlier section) to test for the presence of an ITD effect. The results of these tests are shown in Figure 25. There is a clear trend with ILD: as the right electrode amplitude increased, the right ear stimulus was judged as louder than the left on average.

There is no significant, nor consistent, ITD effect. This is consistent with results from normal hearing subjects in which ITD sensitivity disappears when the cochleotopic positions of excitation are significantly mismatched²². This result indicates that the choice of electrode pair affects the binaural sensitivity and, therefore, this choice should be a consideration in attempts to provide binaural advantages.

5.2.5 Summary

The subject demonstrated a sensitivity to ILD and ITD for trains of pulses (100pps) presented to a pair of electrodes judged to be in similar cochleotopic positions (based on across-ear pitch comparisons). The subject showed sensitivity to ITDs of 300 μ sec and 600 μ sec. No ITD sensitivity was found for an interaural electrode pair that was mismatched cochleotopically. These results are evidence that it may be possible to exploit binaural sensitivity to make binaural advantages available to cochlear implant users. The results also indicate that the choice of electrode pair affects the binaural sensitivity and, therefore, electrode/channel matching across ears should be a consideration in attempts to provide binaural advantages.

6 Acoustic Simulations and New Processor Designs

6.1 Comparison of Speech Reception Performance for Normal and Implanted Listeners

Over the course of this contract, there has been considerable interest in developing acoustic simulations of speech processors used in cochlear implant systems e.g., 25, 26. The input acoustic signal is processed initially as in an implant speech processor (typically, band-pass filtering and envelope detection), but instead of delivering the frequency-specific information from each channel to a different electrode, the channel inputs modulate tonotopically ordered acoustic carriers. The carriers can be either sinusoids or bands of noise. The modulated carriers are summed to form a composite signal which is presented acoustically. The intent of these simulations is to enable normal-hearing listeners to experience an approximation of what an implant subject hears

We utilized acoustic simulations to explore the effects of different numbers of CIS ¹⁴ processing channels in quiet and in noise. The simulations were implemented using a real-time signal processing system based on a Motorola 96000 floating-point DSP with a sample rate of 12 kHz. A block diagram of the simulation is shown in Figure 26. In our experiments: (1) the Preemphasis/AGC was bypassed because it is not used by our pool of implantees; (2) the Nonlinear Mapping was disabled because the small dynamic range associated with electric stimulation is not an issue for normal hearing subjects, (3) the band envelopes modulated tones (or noise-bands) at their respective filter's geometric center frequency (this presents the band energy at the correct cochleotopic place), and (3) the modulated tones were summed for presentation with a speaker or headphone. The number of processor channels varied from 2 to 16 with the carrier frequencies set as shown by the filled circles in Figure 27.

In one set of measurements, a single normal-hearing listener was tested on our 24-initial consonant identification test. Speech-spectrum noise (from the HINT corpus²⁷) was mixed with the test speech at different speech-to-noise ratios (SNRs), at the input to the simulation system, so that both the speech and the noise were processed via the simulation. Measurements were made using simulations with 3, 6, and 12 channels, and with the simulation bypassed -- an unprocessed condition.

The same 24-initial consonant test was also administered to three of our better implant users. Two used a CIS Clarion system (8 channels; 833 pps/electrode) and one, an Ineraid subject, used the MIT version of CIS processing (6 channels; 2000 pps/electrode) implemented on a Geneva wearable sound processor. These subjects were selected on the basis of achieving scores near 80% on the NU-6 monosyllabic word recognition test. Because each of these subjects had worn their CIS sound processor for at least three months at the time of testing, one might be concerned about comparing the implant results with those from the acoustic simulations where the normal-hearing subject did not experience months of continuous listening. While one should be mindful of this difference, we do not believe it is a serious problem because the subject spent

many hours training on these materials (more than 4000 presentations with feedback) to establish a level of asymptotic performance.

6.1.1 Acoustic Simulations: Speech Reception Results

The results plotted in Figure 28 show an orderly decrease in performance as the number of channels and the SNR decrease. In quiet ($\text{SNR} = \infty$), direct listening (Unproc) produced 96% correct, the 12-channel simulation (AS-12) result was about 11 percentage points lower, and each subsequent halving of channels (to AS-6 and AS-3) decreased the performance another 11% percentage points. As SNR decreases, performance falls at varying rates. For intermediate scores, each decrement in the system from unprocessed to AS-3 requires an increase of about 8 dB in SNR to maintain a given level in performance.

The open symbols of Figure 28 represent the mean scores for the three high-performing cochlear implant users described above. Note that in both quiet and in noise, the performance of the implantees is essentially equivalent to that of the normal listener using a 6-channel simulation. This is also evident in Figure 29 where these consonant results are presented in the form of percent information transfer for three consonant features. Notice in the panel for 6-channel processing that the distribution of information transfer across the consonant features for the implantees is very similar to that for the normal-hearing subjects for the noise conditions examined. This is consistent with the view that the implantees and normal hearing listener are extracting similar information from their respective processors.

Figure 29 shows that in quiet ($\text{SNR} = \infty$), the main effect of reducing the number of channels is a large reduction in place information. Since the representation of spectral shape varies directly with the number of processing channels, this effect on place scores is to be expected. In contrast, scores for manner and voicing are affected relatively little, even for the three-channel (AS-3) condition. For decreasing SNR, scores on all features decrease. Manner is generally affected least, place is affected most, and voicing is intermediate. As mentioned above, the average scores from the three high-performing implantees show confusion patterns that are very similar to those of the normal listener using a 6-channel simulation.

In another set of measurements, two groups of four normal-hearing listeners were tested on sentence reception in noise using the HINT test. One group was tested in the unprocessed condition and with simulations using 12, 6, and 3 channels. The second group was tested with simulations using 16, 8, 4, and 2 channels, and at one SNR for unprocessed listening.¹ The results (Figure 30) show that performance falls with decreasing channel number and decreasing SNR. The results are quite orderly within

¹ Two subject groups were required in order to avoid repeating a sentence list during the testing of any individual subject (the limited number of HINT sentence lists was not sufficient for the number of channel/SNR conditions).

each group (Group-1: open symbols, dashed lines; Group-2: filled symbols, solid lines). Some “crossovers” occur between the groups, e.g., Group-2 with 8 channels did better than Group-1 with 12 channels. This may reflect a real difference between the groups and/or variability intrinsic to the test materials.

The three high-performing implantees were also tested on the HINT materials. Figure 31 shows their results along with those from the normals using simulations with a corresponding number of channels (6 and 8). Again, both sets of scores are quite similar. These results also show that for the best performing implantees (with 6 to 8 channels) to attain the same word scores as normals listening without processing (i.e., listening naturally), the implantees require about a 6 dB higher SNR. In some ways, this deficit is remarkably small. It is much less than the deficits obtained with earlier multichannel implants (4-channel Ineraid and feature-based Nucleus systems) and it is similar to that obtained by some listeners with moderate-to-severe hearing impairments using hearing aids. Nevertheless, doubling the effective number of channels (up to 12 or 16) could halve this deficit to about 2 – 3 dB.

6.1.2 Acoustic Simulations: Possible Implications

The main picture is clear: with everyday-like sentences, only a few (perceptually independent) channels are needed to allow high performance in quiet ^{28-30, 25, 31, 26, 32}. However, to allow performance in noise that approaches normal listening, a much higher number of channels is required.

The equivalence in performance between the high-performing implantees and normal-hearing subjects listening to 6- and 8-channel simulations on both sentence and consonant tests implies that these implant subjects are extracting all of the information relevant to speech reception that is available from the implant. This result is impressive and provides a challenge to developing improved systems. It suggests that further gains in performance can be obtained only by increasing the number of CIS processor channels/electrodes that can be utilized by subjects or by using existing electrodes to deliver additional information beyond that represented by the standard band-envelope signals.

6.2 Increasing the Information Provided by CIS Processors

Since we cannot easily effect an increase in the number of independent electrodes to be stimulated from a speech processor, we have considered increasing the information provided to each of the presently implanted electrodes in the Ineraid implant. In other words, we have attempted to decrease the information discarded when each channel’s bandpass signal is represented only by its envelope. Consider, for example, the output of a band-pass filter (bandpass(t)) that can be represented as the product of an envelope

signal ($env(t)$) and a phase-modulated carrier ($\cos[\phi(t)]$) without any loss of generality; so that

$$bandpass(t) = env(t) * \cos[\phi(t)].$$

If a CIS processing channel accurately extracts the band envelope, $env(t)$ is the portion of the band-pass filter output signal that, after being level-mapped, modulates the pulse train to form a stimulus signal. The phase-modulated carrier signal (also referred to as the excitation signal or the fine temporal structure of the bandpass filtered output signal) represents the information discarded by the CIS processing channel.

One important question is how much information is lost when an N-channel, CIS processor discards the excitation signals $\cos[\phi(t)]_k$, for $1 \leq k \leq N$? If the information in the excitation signals is large relative to the information in the collection of envelopes, considerable performance gains might be realized by restoring some of the information discarded in the excitation signals.

Insight into this issue can be gained from a study where normal-hearing subjects listened to the output of a bank of band-pass filters and speech reception was measured as a function of the number of band-pass filters and the envelope bandwidth³³. In Figure 32 (a reproduction of Figure 3 from the Drullman, et al. Paper), speech reception is plotted as a function of the bandwidth of the band envelope signals with the band-pass filter bandwidth (i.e., number of channels for the fixed frequency range of 100-6400 Hz) as a parameter. Note that the bank of one-octave filters includes six channels and corresponds most closely to the 6- and 8-channel CIS processors and CIS simulations used to collect portions of the data shown earlier.

In the case of an "LP cutoff frequency" of zero, there are no envelope variations, and the only information transmitted to the listener is the sum of the excitation signals from each bandpass channel. Listening to the six-filter system that delivered only the excitation signals, Figure 32 shows that normal-hearing subjects scored over 80% correct on a thirteen-sentence test. This demonstrates that considerable information is included in the 6- or 8-channel excitation signals (i.e., in the fine-time structure of the band-pass filters' output signals) that are discarded in CIS sound processors.

As the filters increase in number (i.e., the bandwidth of the band-pass filters decreases), the excitation signal carries less information. In the case of the 24-channel system (1/4 octave filter bandwidths) with an "LP cutoff frequency" of zero (excitation-only system), normal hearing listeners score only a few percent correct on the sentence test. Indicating that, in this 24 channel case, most of the speech information is in the envelope signals.

Another way to compare the spectral information in a bandpass filter's output signal to the information in the envelope signal is to consider their respective waveforms and spectra for a specific example. The time waveforms displayed in Figure 33 represent a steady-state segment from the vowel portion of a synthesized /da/ that has been

processed in four different ways. In each case the /da/ has been processed by a band-pass filter with -3 dB cutoff frequencies of approximately 360 and 640 Hz (this corresponds to Channel 2 of a 6-channel processor of overall bandwidth 200-6500 Hz). Moving from the bottom panel to the top panel the waveforms are: the output of the band-pass filter (BPFO); the half-wave-rectified output of the band-pass filter (HWR); the full-wave-rectified output of the band-pass filter (FWR), and the quadrature envelope of the band-pass filter output ($env(t)$). Figure 34 displays the corresponding log-magnitude spectra in dB versus frequency in Hz for each of the waveforms of Figure 33.

Notice that the spectral information for the output waveform of the band-pass filter between 400 and 700 Hz is well represented by the spectrum of the half-wave-rectified (HWR) case. Below 400 Hz, the HWR spectrum also includes spectral detail associated with the envelope ($env(t)$, top panel) and above 700 Hz the energy from the harmonic distortion of the rectification process is evident.

In the FWR case, the $env(t)$ spectrum is better represented than when half-wave rectification is used. Like the HWR case, the distortion products related to the rectification process are represented above 700 Hz. Note that the spectral details associated with the information carried in the output waveform of the band-pass filter is degraded by the combination of the $env(t)$ and harmonic-distortion components in the range from 300 to 800 Hz. This means there is little information describing the original waveform's fine structure (excitation).

Since contemporary CIS processor implementations tend to use either a quadrature or FWR operation to generate the envelope of the band-pass filter's output waveform (often with low-pass filtering at 400 Hz or lower), it is clear that the details of the waveform's fine structure are discarded. One possible method for increasing the information content within a channel would be to restore some or all of the information contained in the excitation signal in a way that the brain can interpret. In the following section we describe preliminary results from the first in a series of experiments investigating methods of restoring some portion of the excitation signal to the individual channels of CIS processing.

6.2.1 Hybrid (2HWR/4QUAD) CIS Processor

Figure 34 suggests that simply using HWR without low-pass filtering would restore the excitation information (plus the HWR distortion products) to the channel's waveform. This section describes a hybrid CIS processing system that combines two such HWR channels (channels 1 and 2) with four regular CIS processing channels (channels 3-6). This hybrid system is used to test whether subjects are able to extract more information when the excitation signal (and the HWR distortion products) is included with the $env(t)$ in the modulator signal for the two lowest-frequency channels.

6.2.2 Implementation of the Hybrid Processor

We modified the two lowest channels of a CIS processor configuration as shown in Figure 35. The top block diagram shows a standard CIS operation that employs quadrature detection to compute the envelope of the band-pass filter output ($env(t)$). The $env(t)$ is then “compressed” by a level-mapping function and used to amplitude modulate a biphasic pulse train (~ 4000 Hz). The modified implementation for channels 1 and 2 (shown in the lower block diagram of Figure 35) uses the same band-pass filters as our standard CIS processor, but the modulator signal is produced with HWR without low-pass filtering.

Since the two HWR channels carry information with bandwidths extending to approximately 1400 Hz, we increased the carrier rate of these channels to approximately 8 kHz. The carrier modulated by the level-mapped $env(t)$ in channels 3 through 6 was approximately 4 kHz. The order of interleaving for the carrier pulses is shown in the bottom panel of Figure 35. This interleaved pulsatile sequence provides the non-overlapping pulse trains associated with CIS processors, while providing such pulse trains at twice the update rate for channels one and two.

6.2.3 Psychophysical Procedures Used for Fitting the Hybrid Processor

For each electrode, psychophysical measures of threshold (THR) and the most comfortable stimulus level (MCL), were made using a 300 ms duration, unmodulated segment of the carrier used in that electrode’s channel (e.g., the repetition rate for electrodes 3-6 was approximately 4000 pps and for electrodes 1 and 2 approximately 8000 pps).

The hybrid processor was designed by first implementing a standard, 6-channel CIS processor *see 10, 1* using 4000 pps carriers for channels 3-6, 8000 pps carriers for channels 1 and 2, and quadrature-derived envelopes across all channels. After verifying that the standard system performed as expected (e.g. speech reception equal to the subject’s wearable CIS system), channels one and two were switched to HWR operation. Each of the two HWR channels were then balanced for loudness with their corresponding standard quadrature channel by adjusting the I_{max} parameter of each HWR channel’s level-mapping function to give the same loudness while listening to the Iowa, 16-consonant review list ²⁰. After fitting five subjects using this method, we found that I_{max} for the HWR channels was increased by a factor of 1.15-1.20 over a standard CIS implementation.

6.2.4 Comparison Data for CIS and 2HWR/4QUAD Hybrid Processors

The data of Table VII present consonant and vowel identification scores for five Ineraid-implanted subjects using the standard CIS and the 2HWR/4QUAD hybrid sound processors. The consonant scores represent percent correct identification using the Iowa,

16-consonant test (/aCa/, e.g., "asha," "apa") and the Iowa, 8-vowel test (/hVd/, e.g., "heed", "had"). Each of the consonant scores represents 5-20 presentations of the 16-consonant set and each vowel score is derived from 6-12 presentations of the 8-vowel set.

Table VII. Speech Reception --Standard CIS vs Hybrid Processor
[percent correct (standard deviation)]

Subject	16-Consonant Test		8-Vowel Test	
	Standard CIS	Hybrid CIS	Standard CIS	Hybrid CIS
A01	69% (4.3)	73% (3.0)	81% (6.0)	92% (0)
S02	72% (3.8)	77% (2.1)		
S11	63% (3.2)	54% (2.2)	69% (3.6)	67% (2.4)
S18	51% (3.2)	56% (2.6)	67% (2.4)	65% (1.7)
S27	61% (3.4)	59% (2.9)	46% (0)	61% (4.5)

These data do not show large, consistent differences in speech reception between the two processors. In the case of consonant recognition, the only significant difference (.01 level) between the standard and hybrid systems is the lower performance exhibited by subject S11 with the hybrid system. At the same time, two subjects, A01 and S27, posted vowel scores that showed a significant improvement for the hybrid system. One might expect the additional information in the HWR channels to enhance vowel discrimination. In order to understand how our subjects use fine-timing (excitation) information we need to shift our focus to psychophysical measurements of frequency discrimination for our group of Ineraid-implemented subjects as well as one or more normal hearing subjects.

Figure Captions

Figure 1. Components of the wearable sound processing system. Starting at the bottom left: sound processor, processor/earhook cable and connectors, earhook assembly, and the earhook/pedestal cable and connector.

Figure 2. Schematic of two Draper Laboratory voltage-to-current source (V/I) converters. The figure is divided into three segments by the vertical dashed lines. See text for explanation.

Figure 3. Layout of the subassembly that includes the power condition unit, three V/I units and the D/A unit. Each of the individual units uses multi-chip module (MCM) technology to increase the circuit density. The overall size of the subassembly is approximately 4.5 x 8.5 cm.

Figure 4. The redesigned analog input circuit for the Geneva Wearable Processor (GWP). The two audio input paths allow for both earhook microphone and external sound inputs. Because the Delta-Sigma type of A/D converter includes an anti-aliasing filtering, there is a saving in external filter circuitry as well as an increased operating flexibility compared to the present analog input structure.

Figure 5. This figure describes the laboratory sound processing system that is used for all psychophysical testing as well as speech processor development work. The upper right hand corner of the figure shows the present dual 8 channel 16-bit D/A outputs. These outputs can be used with a high speed sequencer for generation of pulsatile stimuli, or can drive isolated current sources directly. The portion presently under design is shown as the waveform generator and the V/I converter, and is modular so that single or multiple channels can be added as stimulation needs arise.

Figure 6. A detailed schematic of the high rate stimulation system (HRSS). The specification text file that is used to describe the desired output waveform is shown on the lower right hand portion of the figure. As shown, this waveform is scaled in amplitude by real-time inputs from the Dual Floating Point DSP system and then passed to the isolated 16-bit D/A converter and V/I circuit.

Figure 7. Diagram of the CIS channel processing used by MIT. The diagram is drawn to show the non-linear mapping function explicitly. The total gain before the mapper is characterized by a single input gain ("G_{in}"), that results in a signal at the mapper input ("e(t)"), representing the bandpass envelope. The mapper output ("m(t)"), amplitude modulates the pulse train that is converted to a current stimulus ("i(t)"), by the voltage-to-current source converter. ("V/I"). The amplitude of the current pulses is determined by "m(t)" and "G_{out}."

Figure 8. The frequency response of the highpass filters used in the front end processing for the Geneva and RTI implementations of CIS running on the GWP. The Geneva filter is implemented with analog hardware, while the RTI filter is implemented in software. The responses are indistinguishable.

Figure 9. The frequency responses for the bandpass filters used in the MIT, Geneva, and RTI processors implemented in the GWP. The Geneva and MIT designs use finite-impulse-response designs, while RTI uses infinite-impulse-response designs. For MIT and Geneva, the lower filters (1-3) are computed at a lower rate than the upper filters (4-6).

Figure 10. The frequency responses for the lowpass smoothing filters used in the three sites' CIS processors. For MIT and Geneva, two designs are required since the envelopes are computed at two rates. A single design is used in the RTI processor.

Figure 11. A set of mapping functions demonstrating how the low and high-level boundaries of the mapping function's inputs were varied to study the effects of input range variations. The boundaries are given for channel #2 in terms of its cumulative level distribution (see Figure 12).

Figure 12. Level histograms and cumulative level distributions for envelope signals in each channel of our standard CIS processor using the standard TIMIT database as input. These statistics were computed on the envelope signals before G_{in} of the mapping stage (see Figure 7).

Figure 13. LG (magnitude estimate assigned to the different stimulation currents) of electrodes 3 and 6 for subject S04. On this graph we describe the loudness assigned to the I_{max} currents matched to 80 % of the dynamic range.

Figure 14. Speech reception performance obtained with α (white) and L (gray) logarithmic mapping functions. On the left panel the consonant identification scores are from at least 10 randomized presentations of the 24 consonants (or 16 consonants for subject S02, S05, S15). On the right panel the monosyllabic word identifications scores are computed from the presentation of 1 or 2 lists of 50 monosyllabic words. The scores are expressed in percent correct. Error bars represent standard deviations.

Figure 15. Subject S01's most apical electrode's loudness growth function. The solid line represents the 4th order polynomial fitted loudness growth function $L_e(I_e)$. The error bars represent the interquartiles of the 8 magnitude estimates obtained for each of the 20 different amplitudes evenly distributed from THR and MCL; the means are represented by stars. The two dotted lines represent the 95% confidence limits for the fitted LG. R^2 is the multiple correlation coefficient.

Figure 16. Loudness Growth functions obtained for the seven subjects for each of their electrodes. The solid lines represent the 4th order polynomial fitted loudness growths $L_e(I_e)$ defined from magnitude estimates realized between the THR and the MCL for each electrode.

Figure 17. Loudness Growth functions obtained for each electrode of subject S01 (left panel) and of subject S15 (right panel). The solid lines are the fits to the AME results and the dashed lines are examples of exponential growth functions over the same range.

Figure 18. NLG mapping functions designed to restore normal LG for single tones, for each electrode for all seven subjects. The diagonal lines represent logarithmic mapping functions.

Figure 19. The left panel shows consonant identification performance for CIS processors using L logarithmic mapping functions (gray) and NLG mapping functions (white). Mean consonant identification scores are computed using at least 10 randomized presentations of the consonant lists for each subject. The right panel presents results for single-syllable word recognition.

Figure 20. Data from the Geneva group comparing CA and CIS scores for 22 Ineraid-implemented subjects for a 14-item consonant test. The CA processor used was the commercial Ineraid system.

Figure 21. Data from MIT comparing CA and CIS performance for 18 Ineraid-implemented subjects for a 16-item consonant test. The top panel presents data obtained on the day that subjects were fit with the CIS system running on the GWP. The bottom panel presents data obtained after six months of CIS use. The CA scores used were obtained with the subjects' Ineraid processors (at least six months experience).

Figure 22. Data comparing mean scores for CA and CIS processor-use at Geneva and MIT. The Geneva scores are based upon 14-item consonant tests, the MIT scores used 16-item consonant tests.

Figure 23. This schematic shows the first two pulse pairs in a bilateral train of pulses. Stimuli delivered to the right electrode (Ineraid) are always delayed relative to those delivered to the left electrode (Clarion). The time axis is to scale and the ITD shown is 600 μ sec.

Figure 24. Lateral position as a function of ITD and ILD. This figure shows the mean (\pm standard error) lateral position for the indicated ILD and ITD conditions. Electrode 7MC was held at 551.8 μ App while the amplitude of 2I was varied. The arrow indicates the 2I stimulus amplitude measured during the centering test. Non-zero ITD data are shown slightly offset in amplitude for easier visual comparison, although their stimuli were presented at the same amplitudes as the zero ITD stimuli.

Figure 25. Relative Loudness as a function of ITD and ILD. The figure shows the mean (\pm standard error) relative loudness for 5 ILDs and 3 ITDs. The electrode 3MC amplitude was held at 408.8 μ App” while the stimulus amplitude of electrode 2I was varied. The arrow indicates the balanced 2I stimulus amplitude measured during the simultaneous loudness balancing test. Non-zero ITD data are shown slightly offset in amplitude for easier visual comparison although their stimuli were presented at the same amplitudes as the zero ITD stimuli.

Figure 26. Block diagram of system used to simulate a CIS processor for normal-hearing listeners. In many CIS systems, preemphasis and/or AGC conditions the input signal. Because the CIS system our implantees use does not include this type of conditioning, it was not used in the simulation. The number of channels (band-pass filters) was a parameter that was varied. The bandwidths for the various conditions are shown in Figure 27. The envelopes were computed using full-wave rectification and a low-pass filter with cut-off frequency of 400 Hz. The nonlinear map used in CIS processors was disabled for these simulations. For each channel, a tone at the channel’s geometric center frequency (see Figure 27) was modulated by its respective envelope. These AM signals were summed and presented to normal-hearing listeners by speaker or headphone.

Figure 27. Cut-off frequencies and “center frequencies” for the individual sets of channels. The numbers at the right specify the number of channels for each channel set. The vertical lines represent the cutoff frequencies of the band-pass filters used to implement each set of channels. The filled circles show the frequency of the sinusoidal carrier associated with each channel (Tone 1, ... Tone n in Figure 26).

Figure 28. Speech reception results for a single, normal-hearing listener (filled symbols; dashed lines) listening to: unprocessed 24-initial consonants, and 12, 6, and 3 channel simulations of implant speech processors whose inputs are the 24-initial consonants. Average results for three high-performing implant subjects using their normal sound processor for the same inputs are also shown (open symbols, solid line). Data are shown for a range of additive noise conditions.

Figure 29. A summary of consonant confusions, in terms of the percentage information transmission for the features: voicing, manner, and place. Data are presented for the various conditions of unprocessed, 12, 6, and 3 channel simulations, and for a range of SNRs at which the 24-consonants are presented. Results for the cochlear implant users (CI) are presented in the same panel as the 6-channel (AS-6) results for the normal-hearing subject listening to the acoustic simulation.

Figure 30. Results from two independent groups of four listeners tested on sentence reception in noise using the HINT test. One group was tested in the unprocessed condition as well as with simulations using 12, 6, and 3 channels. The second group was tested with 16, 8, and 2 channel simulations, and at one SNR for the unprocessed condition.

Figure 31. Individual data for the sentences in noise test (HINT) for the same three high-performing implanted subjects (open symbols, solid lines). Average data from normal subjects without processing (solid symbols, solid line) and normal subjects using the 6 and 8 channel simulations (solid symbols, dashed lines).

Figure 32. Reproduction of Drullman, et.al. Figure 3³³.

Figure 33. The time waveforms (arbitrary amplitude vs. sample number) derived from a steady-state segment from the vowel portion of a synthesized /da/ that has been processed in four different ways. In each case the /da/ has been processed by a band-pass filter with -3 dB cutoff frequencies of approximately 360 and 640 Hz (this corresponds to Channel 2 of a 6-channel processor of overall bandwidth 200-6500 Hz). Moving from the bottom panel to the top panel the waveforms are: the output of the band-pass filter; the half-wave-rectified output of the band-pass filter (HWR); the full-wave-rectified output of the band-pass filter (FWR), and the quadrature envelope of the band-pass filter output (env(t)). The sample interval associated with each waveform 62.5 μ s.

Figure 34. Amplitude spectra of the waveforms shown in Figure 33.

Figure 35. Block diagrams and carrier-timing structure for the two styles of processing used to implement a hybrid CIS processor. The top block diagram represents the processing used by the four highest-frequency channels (channels 3-6). A quadrature method is used to compute an envelope that is "compressed" using a level-mapping function and then serves to amplitude modulate a biphasic pulse train (16 μ s/phase, cathodic first, 3907 pps). The lower block diagram represents the processing used by the two lowest-frequency channels (1 and 2) and shows the modulator being derived by half-wave rectification (HWR) and level mapping. The carrier is a biphasic pulse train like that used for channels three through six with a repetition rate of 7813 pps. The bottom panel shows the ordering of the interleaved pulses across the channels. This interleaved pulsatile sequence provides the non-overlapping pulse trains associated with CIS processors, while providing twice the update rate for channels one and two. For these lower channels, the waveform's zero crossings are defined within $\pm 64 \mu$ s. Obviously the pulsatile carrier rate for channels one and two can be increased if this jitter is shown to degrade the information presented to implanted electrodes.

References

- ¹ Donald K. Eddington et al., "Speech Processors for Auditory Prostheses, First Quarterly Progress Report," MIT Research Laboratory of Electronics, 1996.
- ² Duane Larsen, Donald K Eddington, and Joseph Tierney, "Speech Processors for Auditory Prostheses, Fourth Quarterly Progress Report," MIT Research Laboratory of Electronics, 1996.
- ³ Donald K. Eddington et al., "Speech Processors for Auditory Prostheses, Ninth Quarterly Progress Report," MIT Research Laboratory of Electronics, 1998.
- ⁴ Donald K Eddington, Joseph Tierney, and Victor Noel, "Speech Processors for Auditory Prostheses, Sixth Quarterly Progress Report," MIT, Research Laboratory of Electronics, 1997.
- ⁵ W.M. Fisher et al., The DARPA TIMIT Acoustic-Phonetic continuous Speech Corpus (Palo Alto, California: 1986) 93-9.
- ⁶ Marco Pelizzone and Colette S Boex-Spano, "Personal communication," (1997) .
- ⁷ Donald K. Eddington et al., "Speech Processors for Auditory Prostheses, Second Quarterly Progress Report," MIT Research Laboratory of Electronics, 1996.
- ⁸ Donald K Eddington et al., "Speech Processors for Auditory Prostheses, Seventh Quarterly Progress Report," MIT, Research Laboratory of Electronics, 1997.
- ⁹ D.K. Eddington et al., "Speech Processors for Auditory Prostheses, Seventh Quarterly Progress Report.," MIT, Research Laboratory of Electronics, 1994.
- ¹⁰ Donald K. Eddington et al., "Speech Processors for Auditory Prostheses, Final Report," MIT Research Laboratory of Electronics, 1995.
- ¹¹ H Levitt, "Transformed up-down methods in psychoacoustics," J Acoust Soc Am 49.2 (1971): 467-477.
- ¹² R.S. Tyler, J.P. Preece, and M.W. Lowder, "The Iowa Audiovisual Speech Perception Laser Videodisc," (Dept Otolaryngology Head Neck Surgery: University of Iowa Hospital Clinic, 1987).
- ¹³ F. G. Zeng and R. V. Shannon, "Loudness balance between electric and acoustic stimulation," Hear Res 60.2 (1992): 231-5.
- ¹⁴ B. S. Wilson et al., "Better speech recognition with cochlear implants," Nature 352.6332 (1991): 236-8.
- ¹⁵ Colette Boëx-Spano, Marco Pelizzone, and Pierre Montandon, "Speech recognition with a CIS strategy for the Ineraid multichannel cochlear implant," Am J Otol 17.1 (1996): 61-8.
- ¹⁶ Donald K. Eddington et al., "Speech Processors for Auditory Prostheses, Third Quarterly Progress Report," MIT Research Laboratory of Electronics, 1996.
- ¹⁷ M. F. Dorman and P. C. Loizou, "Changes in speech intelligibility as a function of time and signal processing strategy for an Ineraid patient fitted with continuous interleaved sampling (CIS) processors," Ear Hear 18.2 (1997): 147-55.
- ¹⁸ Blake S Wilson, "Personal communication," (1997).

- 19 Colette S Boëx-Spano, "Codage des sons pour les implants cochléaires," Doctoral Thesis, de l'Universite Joseph Fourier, 1995.
- 20 R.S. Tyler, J.P. Preece, and M.W. Lowder, "The Iowa Cochlear Implant Tests," (Dept of Otolaryngology Head Neck Surgery: University of Iowa Hospital Clinic, 1983) .
- 21 B. S. Wilson et al., "New processing strategies in cochlear implantation," Am J Otol 16.5 (1995): 669-75.
- 22 J M Nuetzel and E R Hafter, "Discrimination of interaural delays in complex waveforms: spectral effects," J Acoust Soc Am 69.4 (1981): 1112-1118.
- 23 N.I. Durlach and H.S. Colburn, "Binaural Phenomena," Handbook of Perception, Vol. 4, eds. E. Carterette and M. Friedman, vol. 4 (Academic Press, 1978).
- 24 J Blauert, Spatial hearing. The psychophysics of human sound localization, Revised ed. (Cambridge: The MIT Press, 1997).
- 25 R.V. Shannon et al., "Speech recognition with primarily temporal cues," Science 270.5234 (1995): 303-4.
- 26 M. F. Dorman, P. C. Loizou, and D. Rainey, "Speech intelligibility as a function of the number of channels of stimulation for signal processors using sine-wave and noise-band outputs," J Acoust Soc Am 102.4 (1997): 2403-11.
- 27 S.D. Soli and M. Nilsson, "Assessment of communication handicap with the HINT," Hearing Instruments 45 (1994): 12-16.
- 28 A. E. Holmes, F.J. Kemker, and G. E. Merwin, "The effects of varying the number of cochlear implant electrodes on speech perception," Am J Otol 8.3 (1987): 240-6.
- 29 M.F. Dorman et al., "Consonant recognition as a function of the number of channels of stimulation by patients who use the Symbion cochlear implant," Ear Hear 10.5 (1989): 288-91.
- 30 D.T. Lawson, B.S. Wilson, and C.C. Finley, "New Processing Strategies for Multichannel Cochlear Protheses.," Natural and Artificial Control of Hearing and Balance, eds. J.A. Allum et al., vol. 97, Progress in Brain Research (Amsterdam: Elsevier, 1993) 313-321.
- 31 Dewey T. Lawson et al., "Speech processors for auditory prostheses, Third Quarterly Progress Report," , NIH # N01-DC-5-2103, 1996.
- 32 K. E. Fishman, R. V. Shannon, and W. H. Slattery, "Speech recognition as a function of the number of electrodes used in the SPEAK cochlear implant speech processor," J Speech Hear Res 40.5 (1997): 1201-15.
- 33 R. Drullman, J.M. Festen, and R. Plomp, "Effect of temporal envelope smearing on speech reception," J Acoust Soc Am 95.2 (1994): 1053-64.

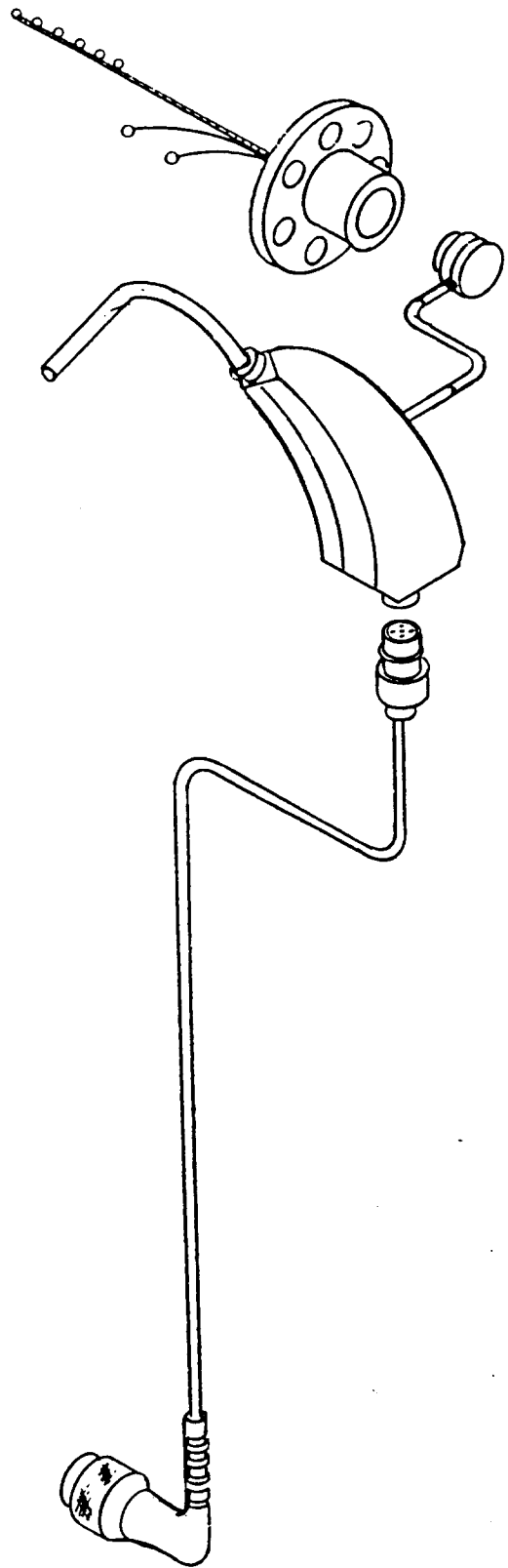
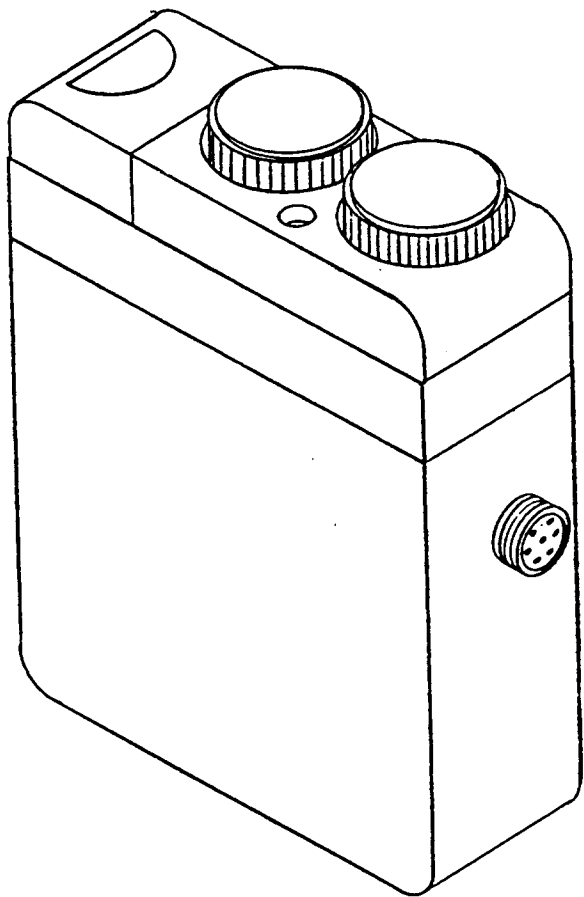


Figure 1

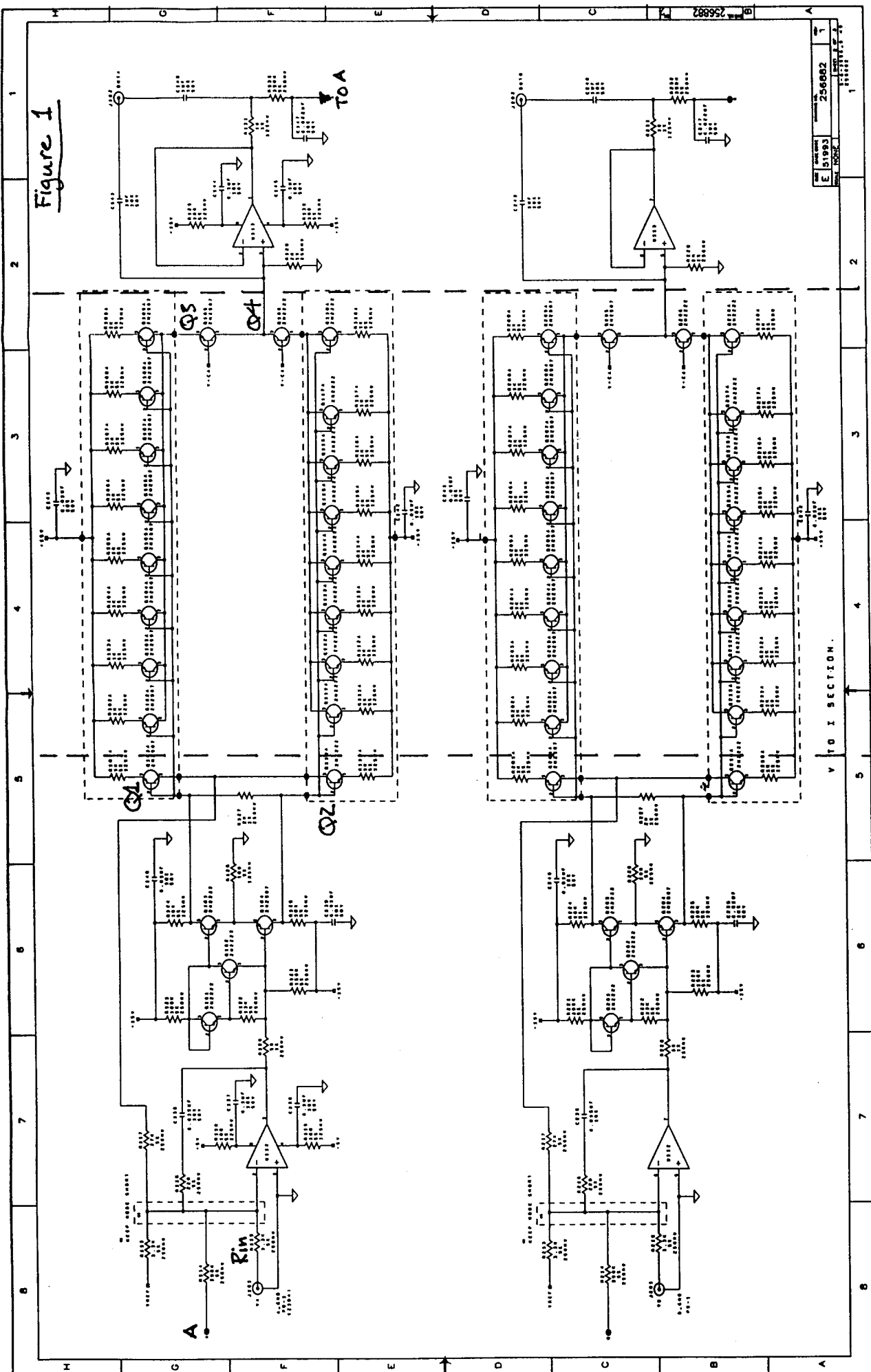
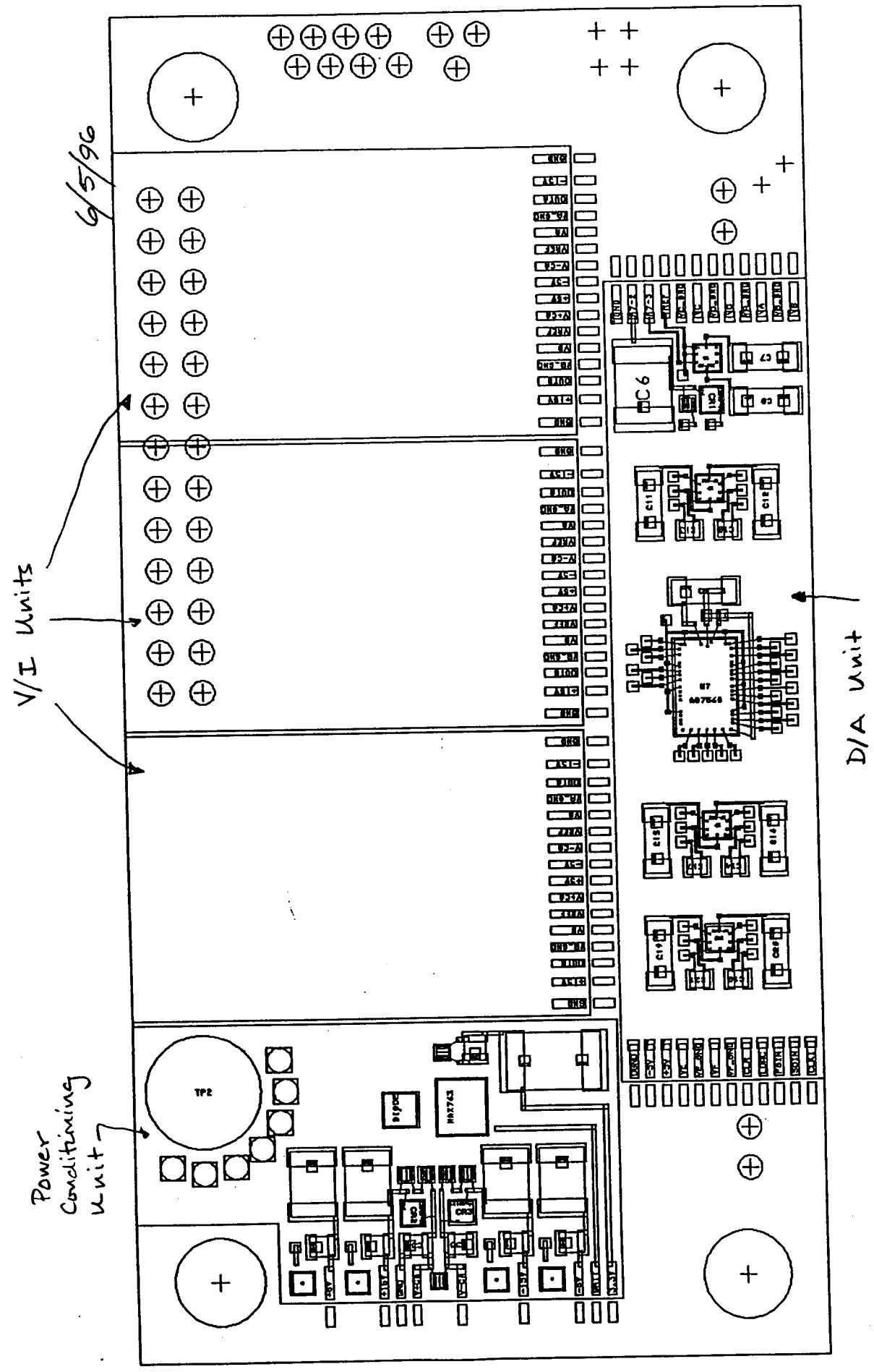
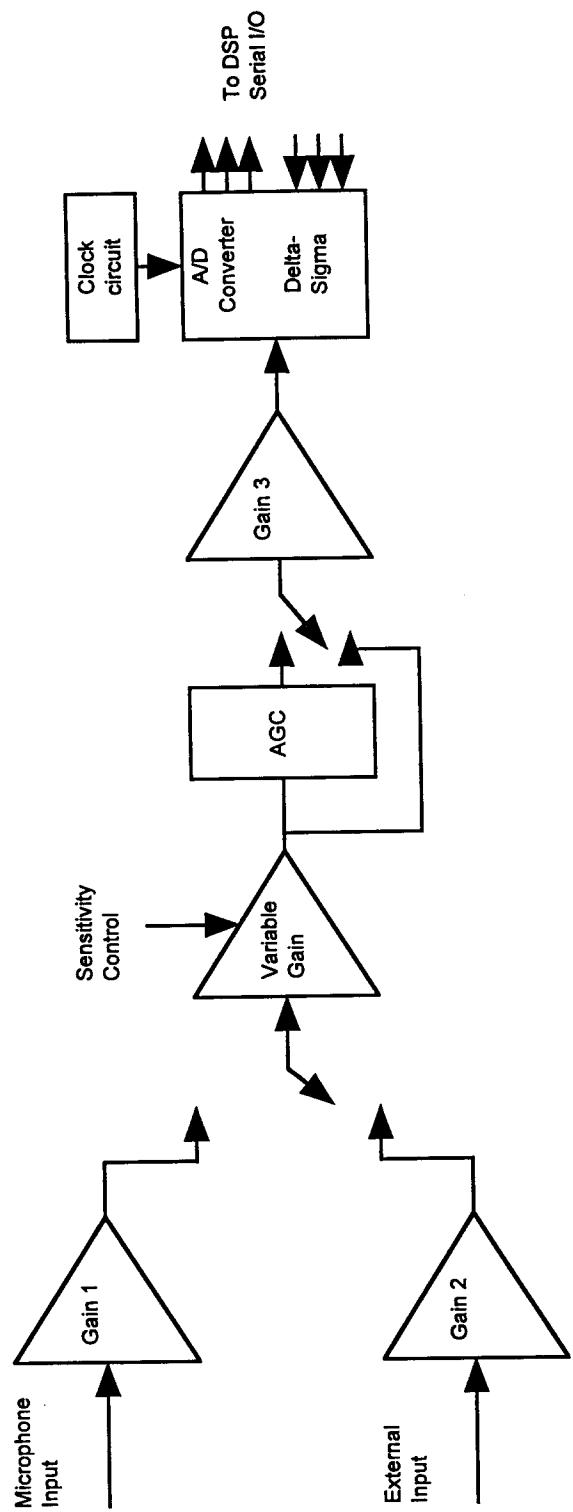
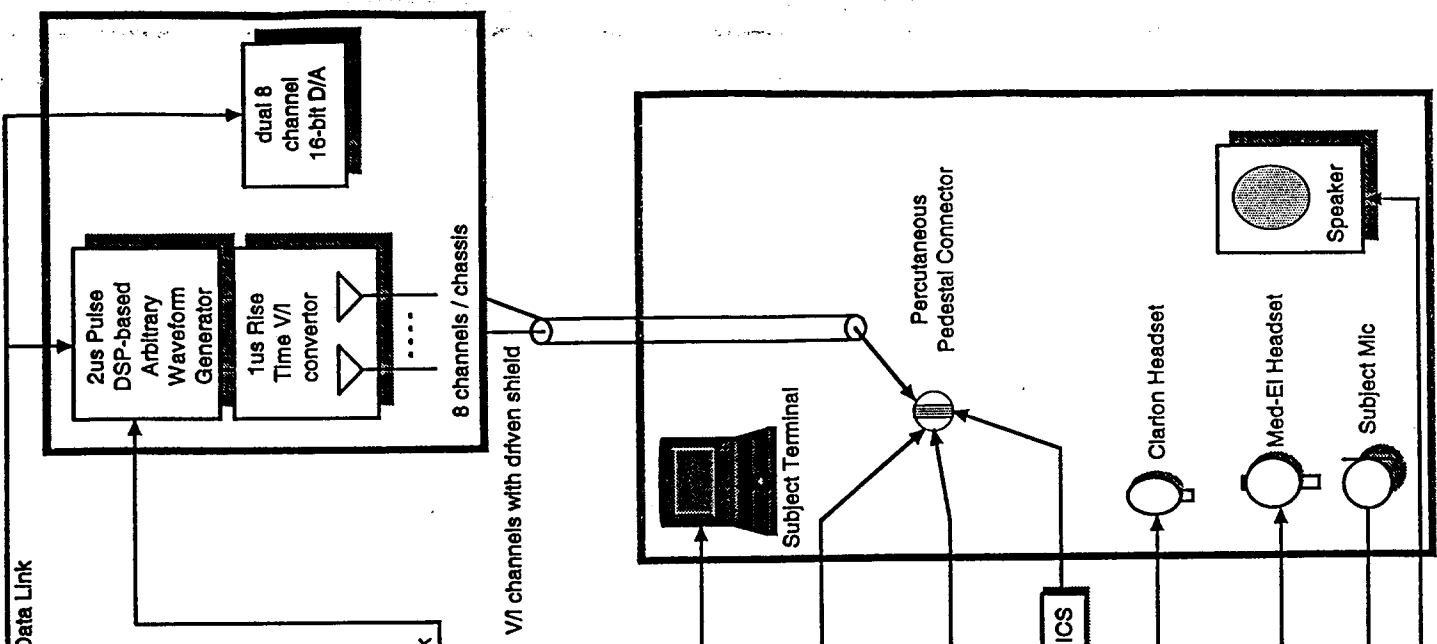


Figure 2

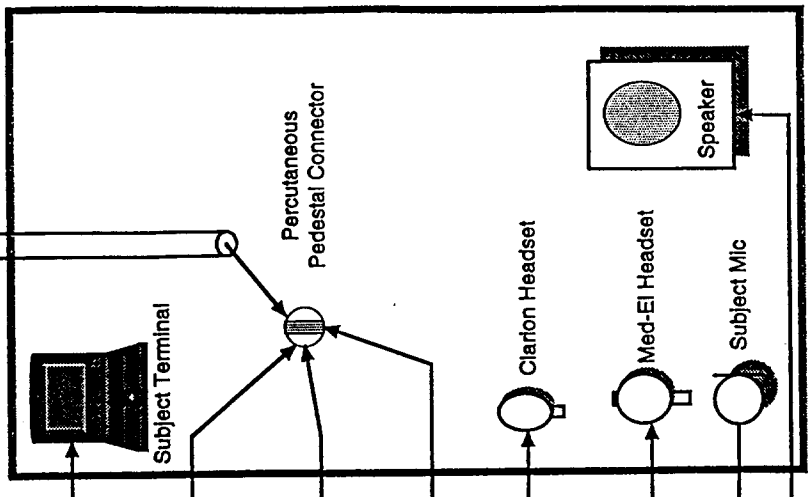




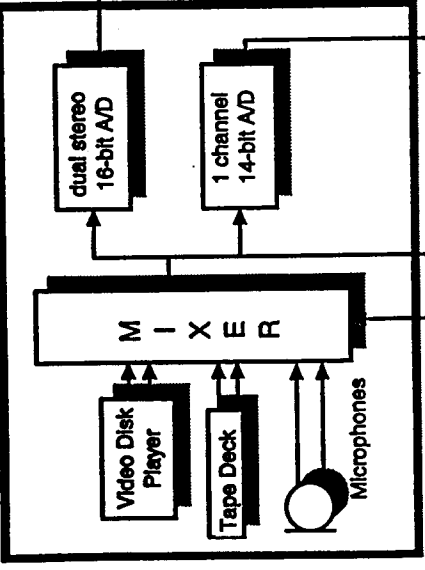
PERCUTANEOUS OUTPUT



SUBJECT SOUND BOOTH



INPUT SUBSYSTEM

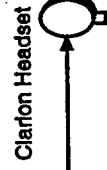
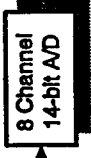
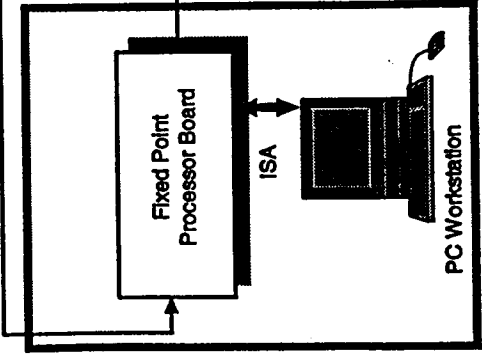
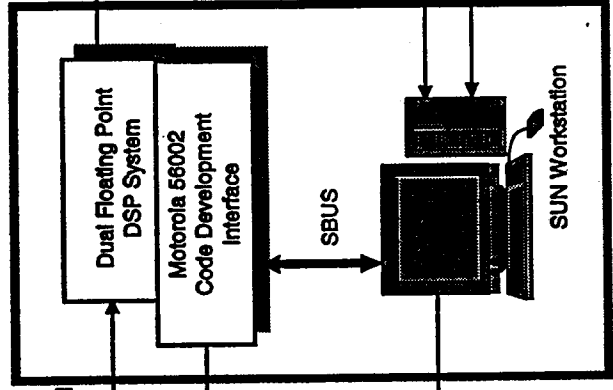


High Speed SSI Data Link

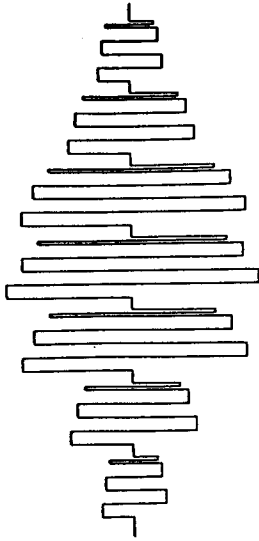
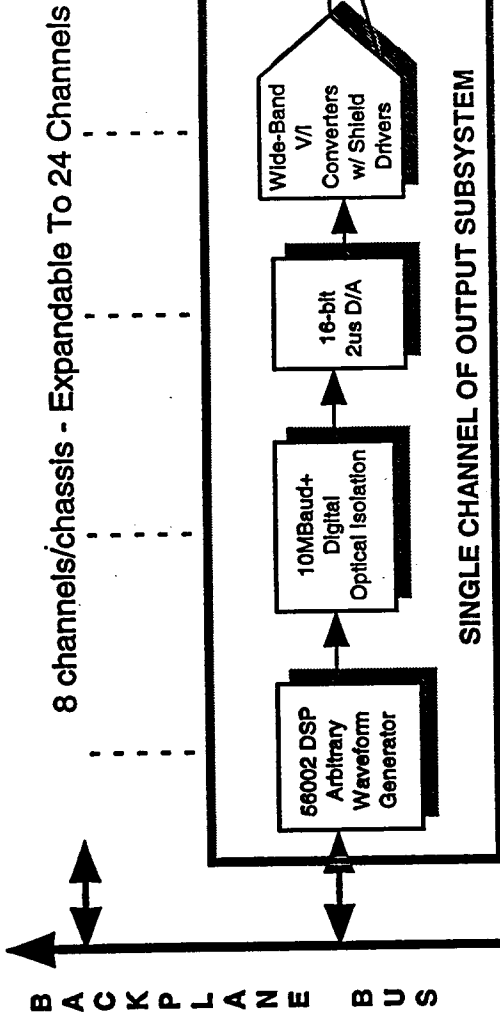
RS232 Program Link

High Speed SSI Data Link

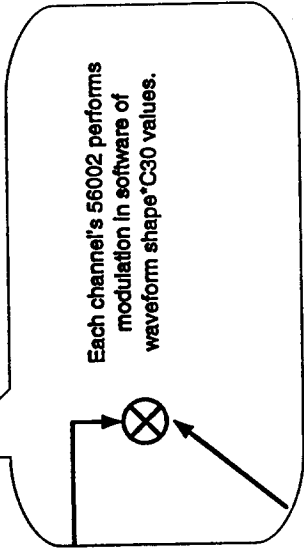
RS232 Link



T O S U B J E C T

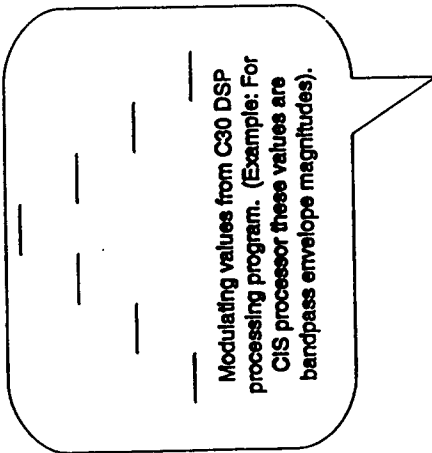
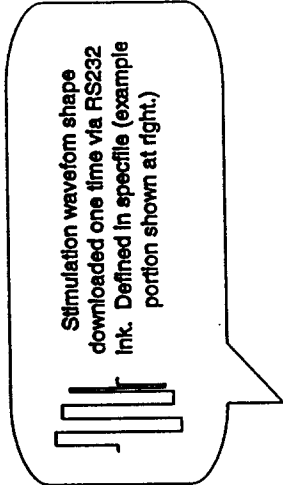


ANALOG OUTPUT CURRENT

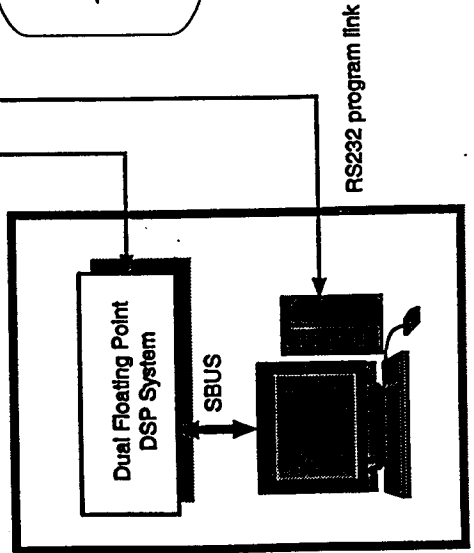


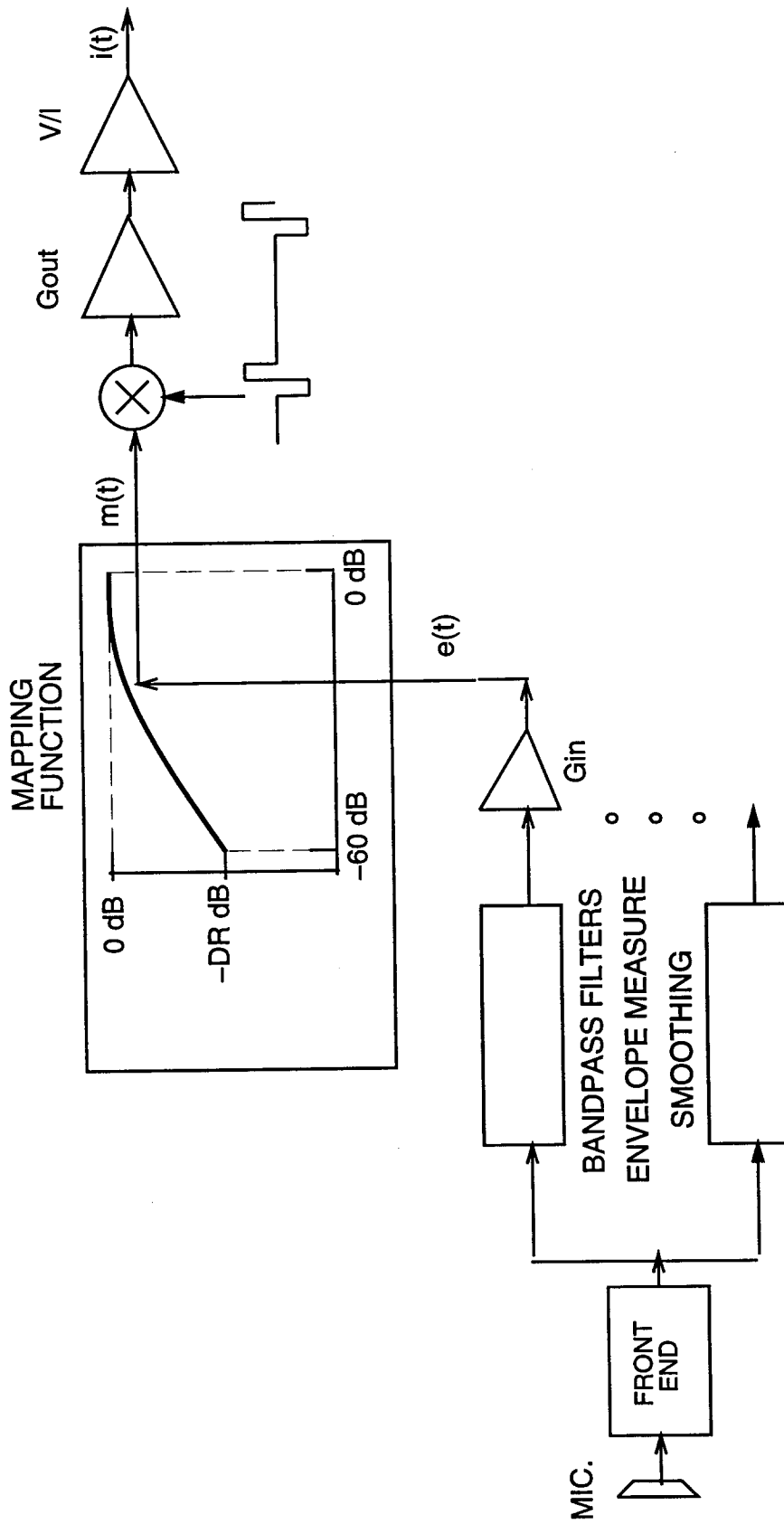
```

"spec file"
channel 4 repeat 0 # repeat forever
.75 4 usecs
-.75 4 usecs
.60 4 usecs
-.60 4 usecs
.5 2 usecs
-.5 2 usecs
end_channel
    
```

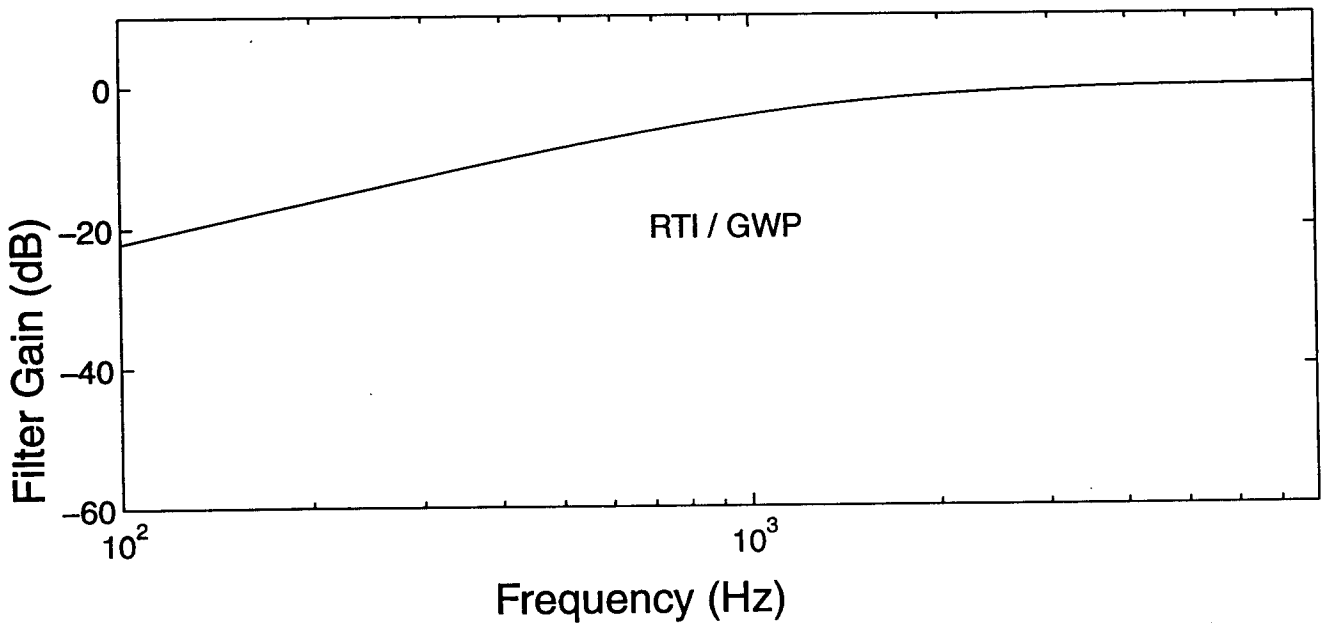
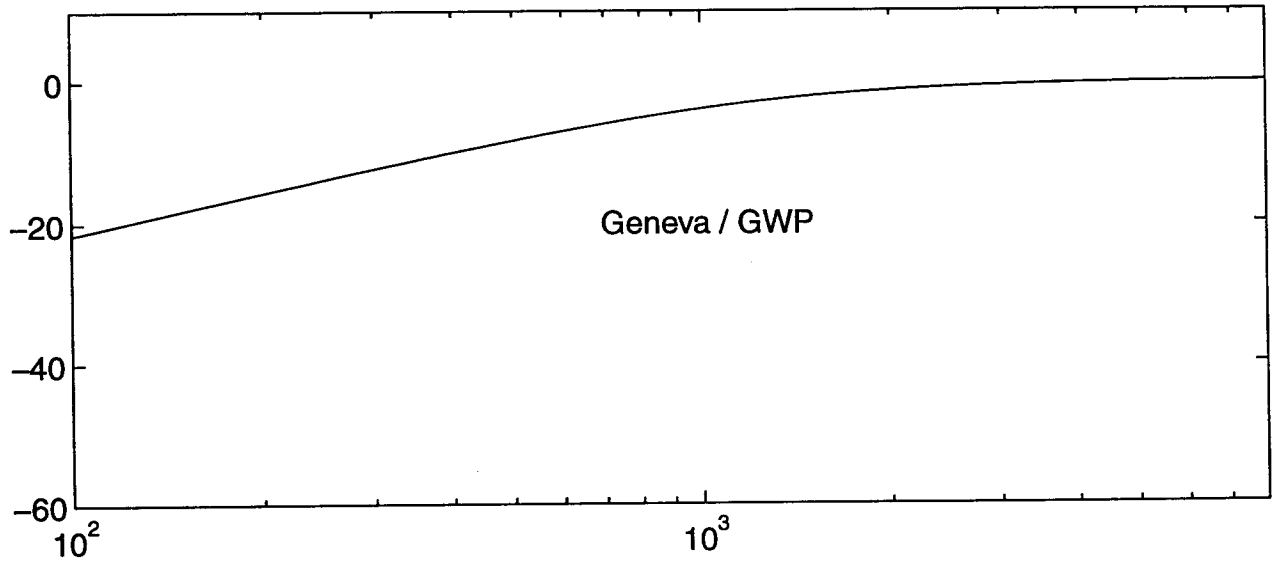


High Speed SSI Data Link

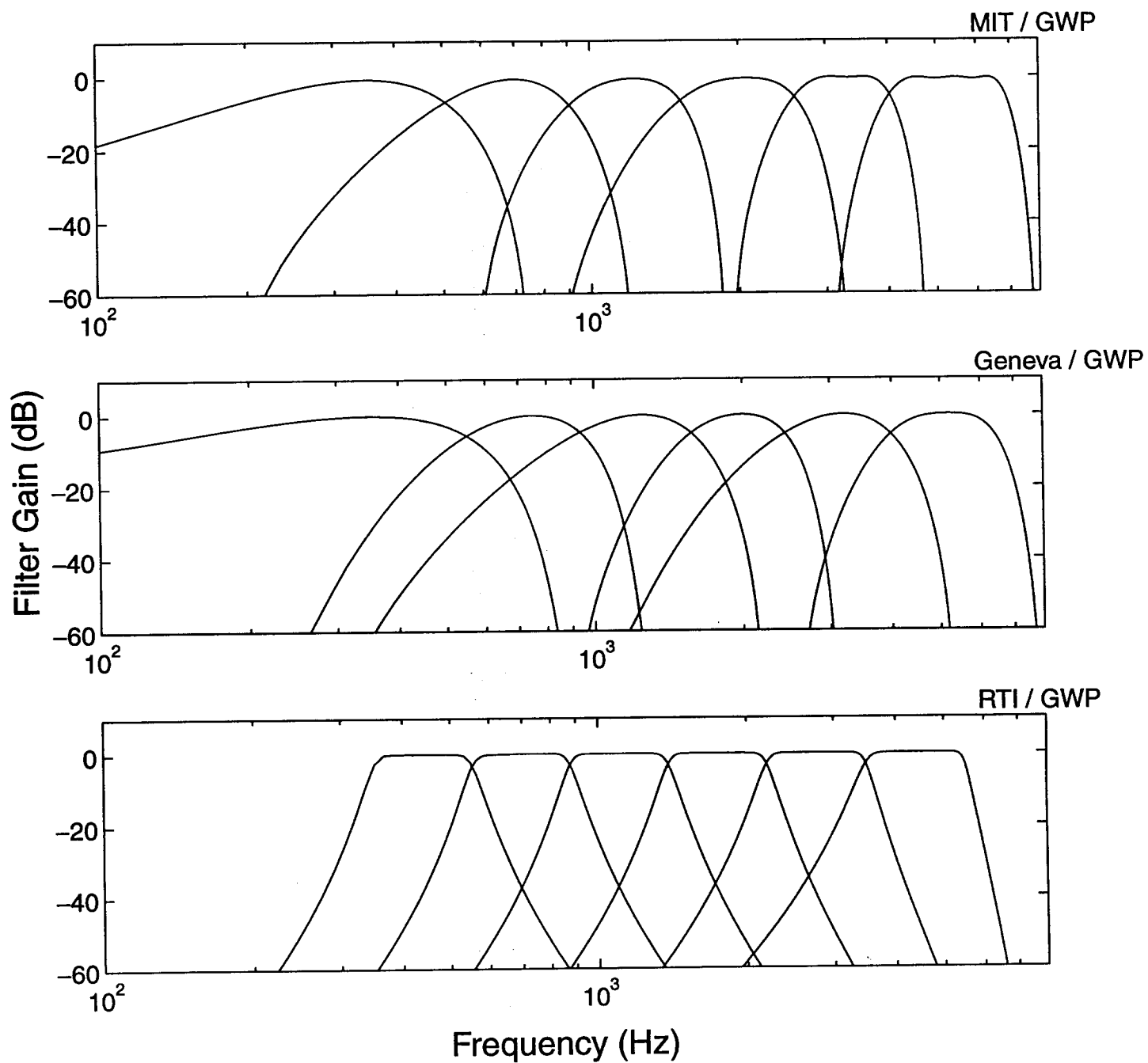




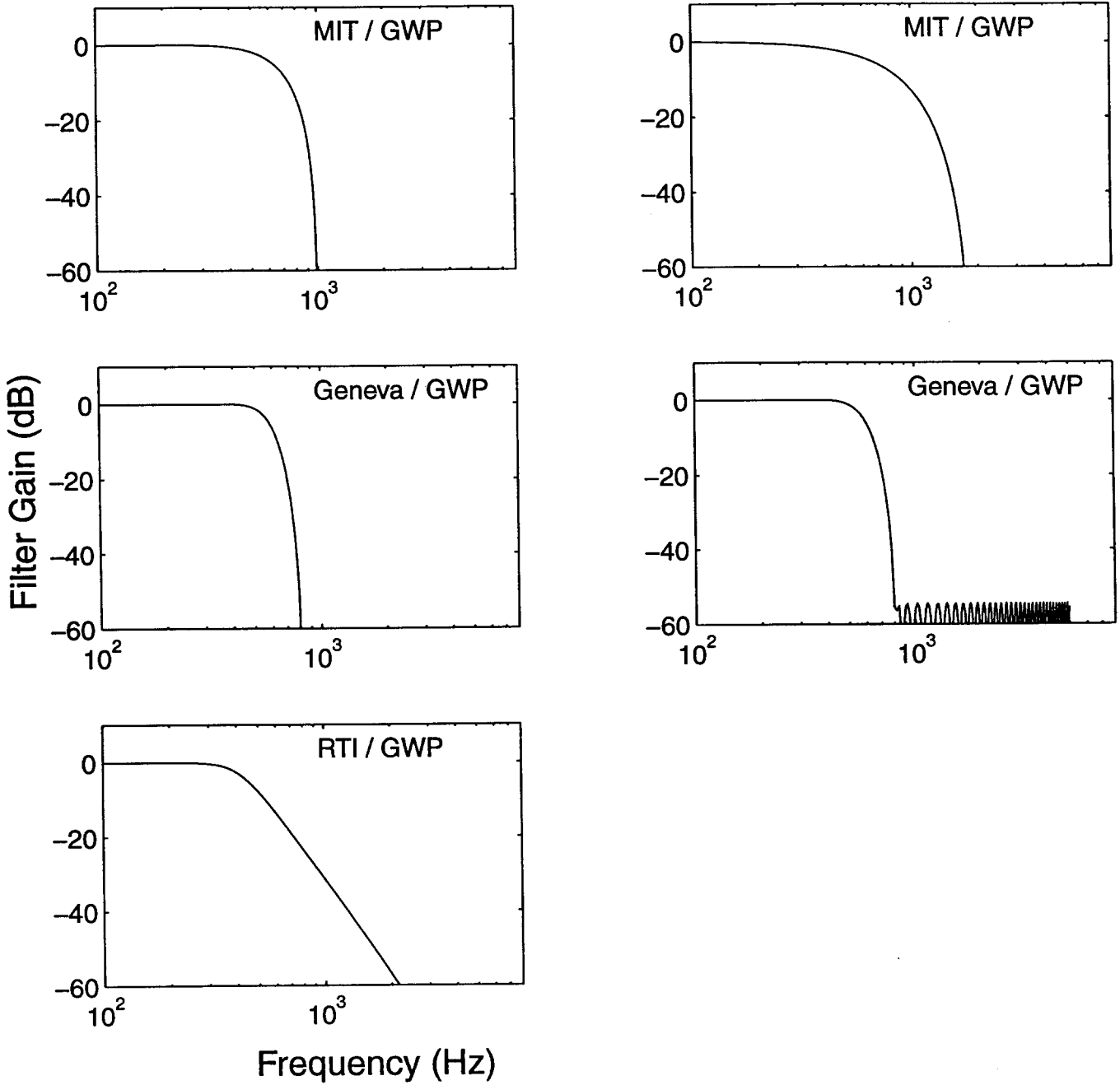
FRONT END HIGH-PASS FILTER RESPONSE



CHANNEL FILTER FREQUENCY RESPONSE

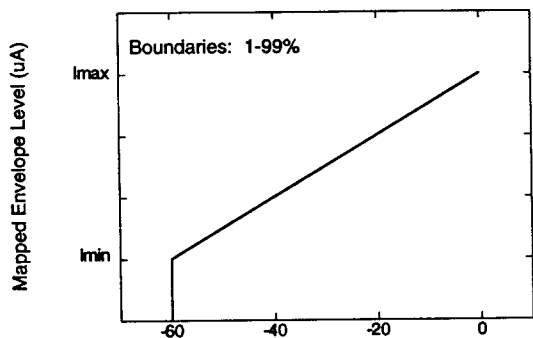


LOW-PASS FILTER FREQUENCY RESPONSE

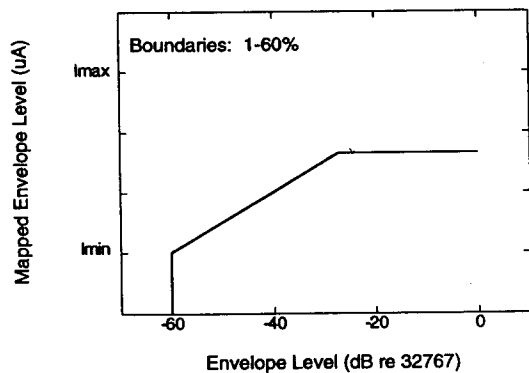
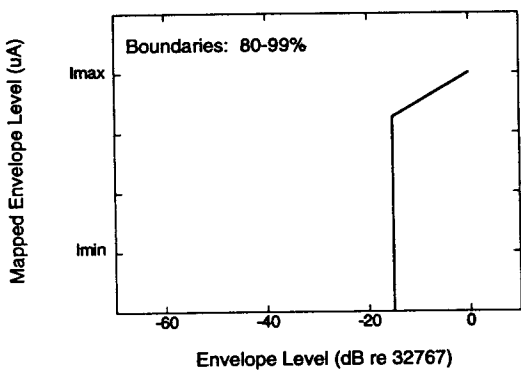
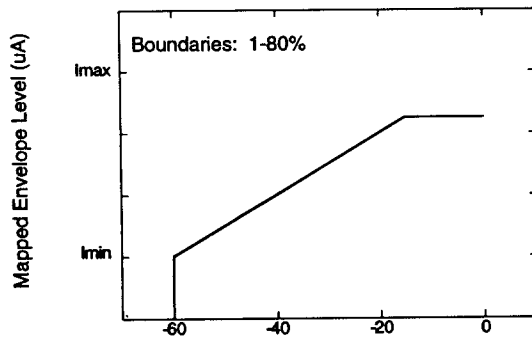
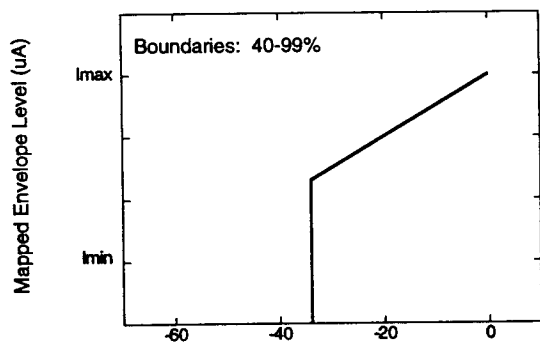
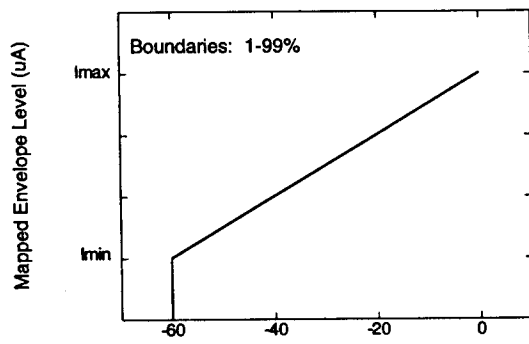


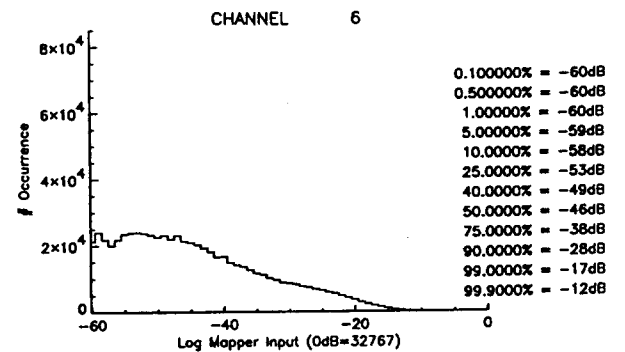
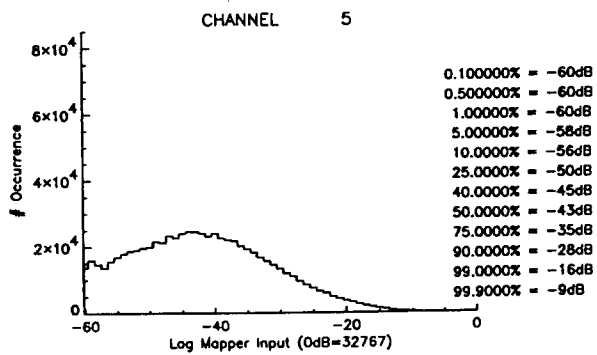
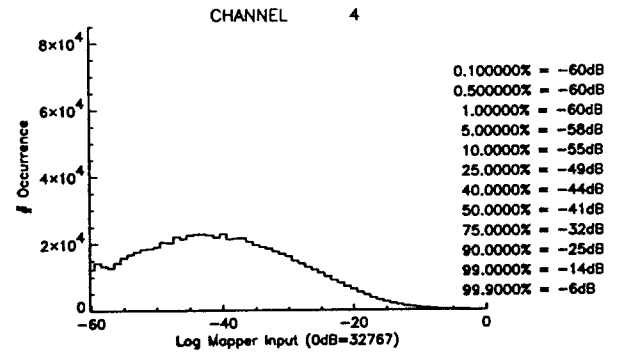
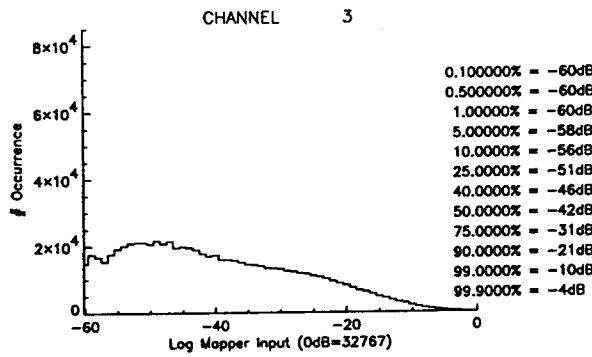
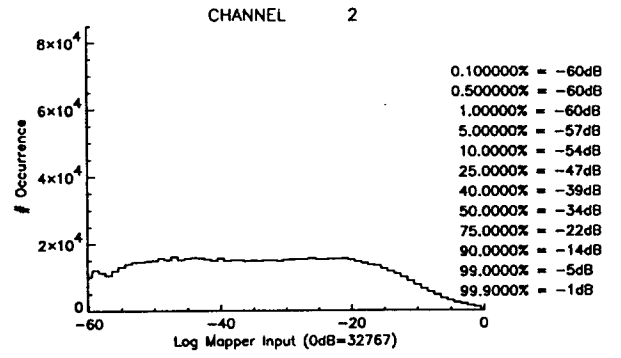
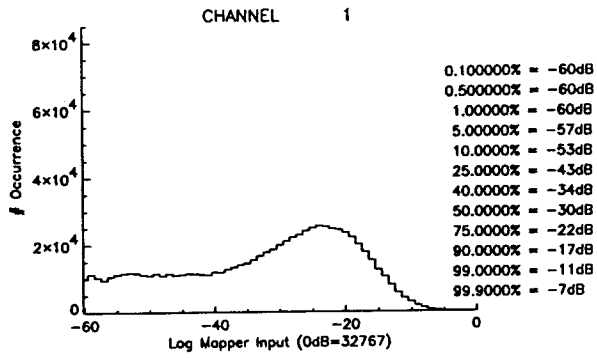
Mapping Range Variations

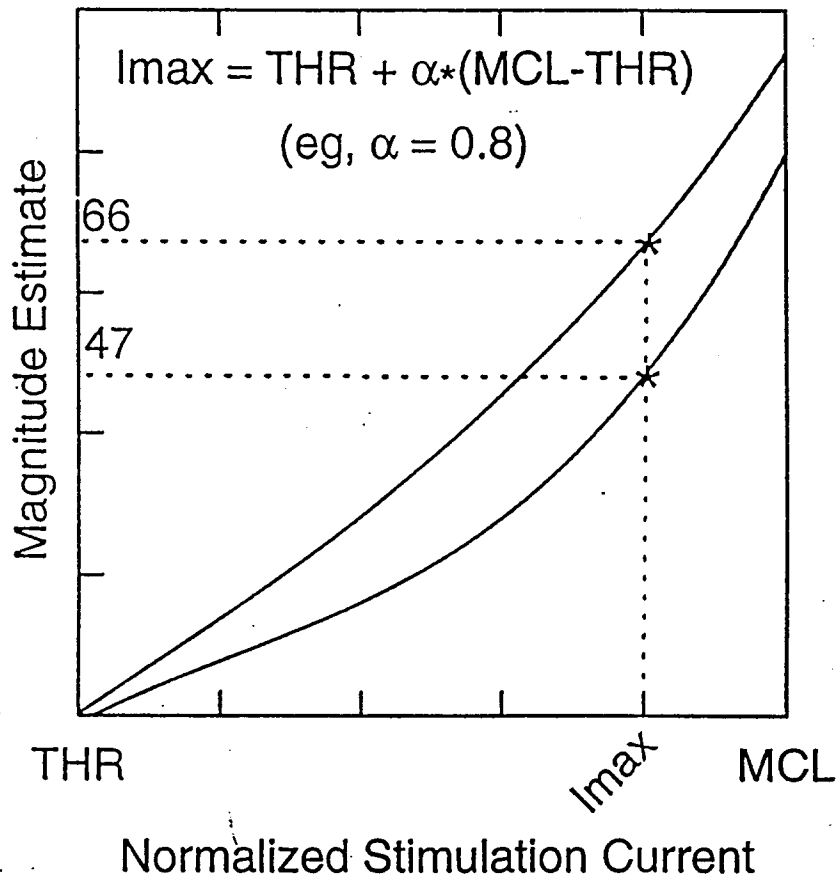
Variation of Low-Level Boundary



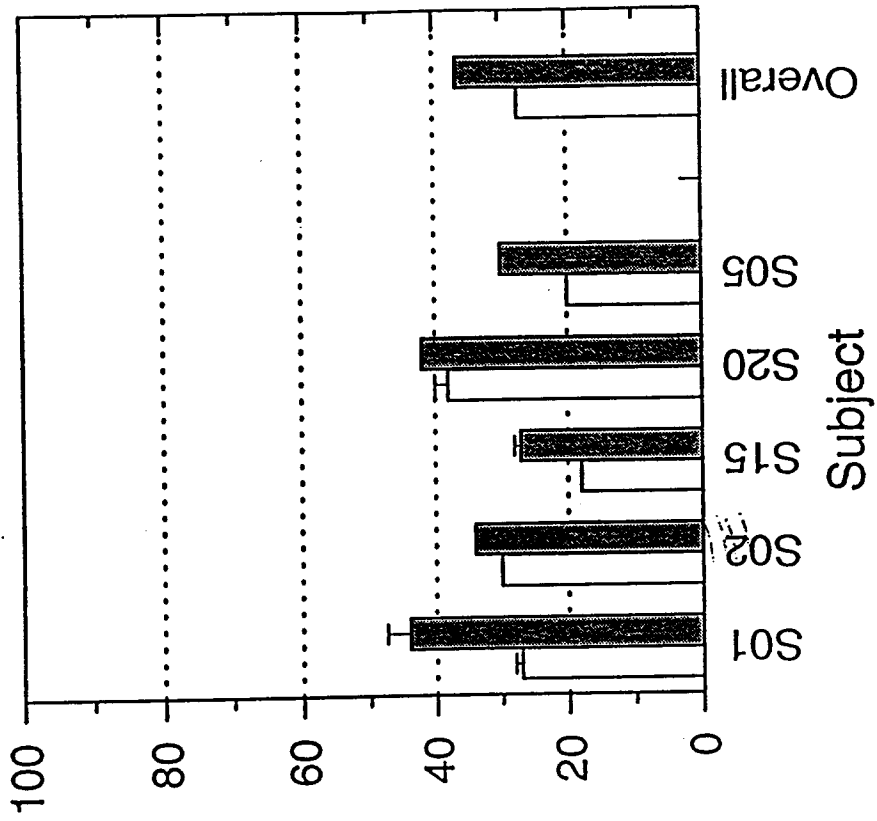
Variation of High-Level Boundary



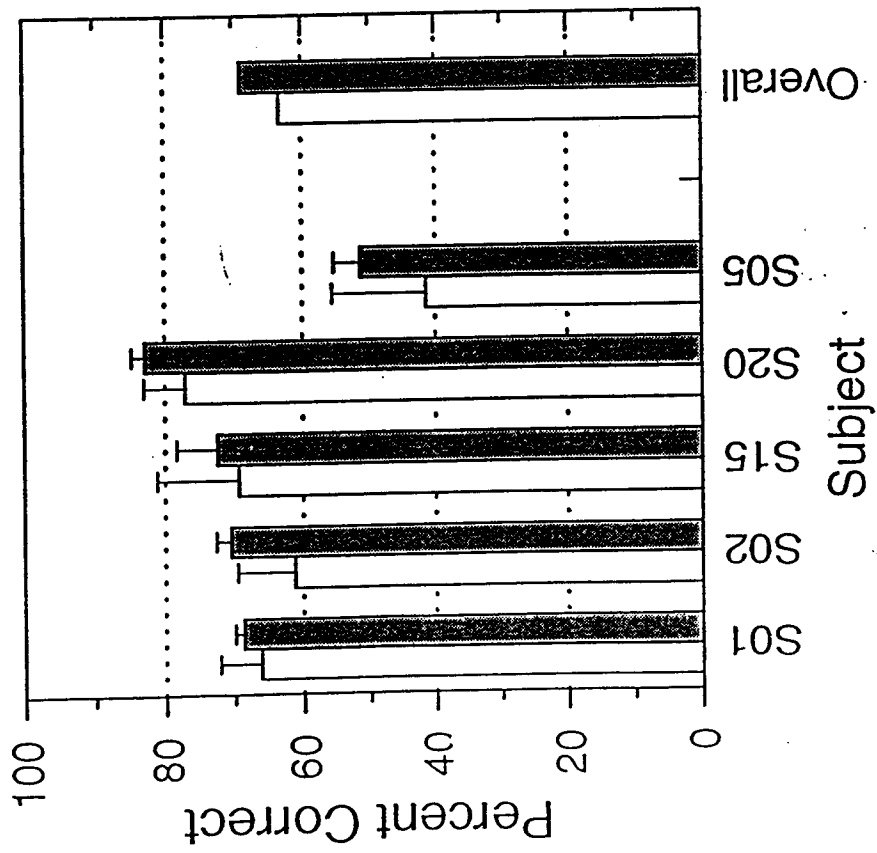






NU-6 Monosyllabic Words

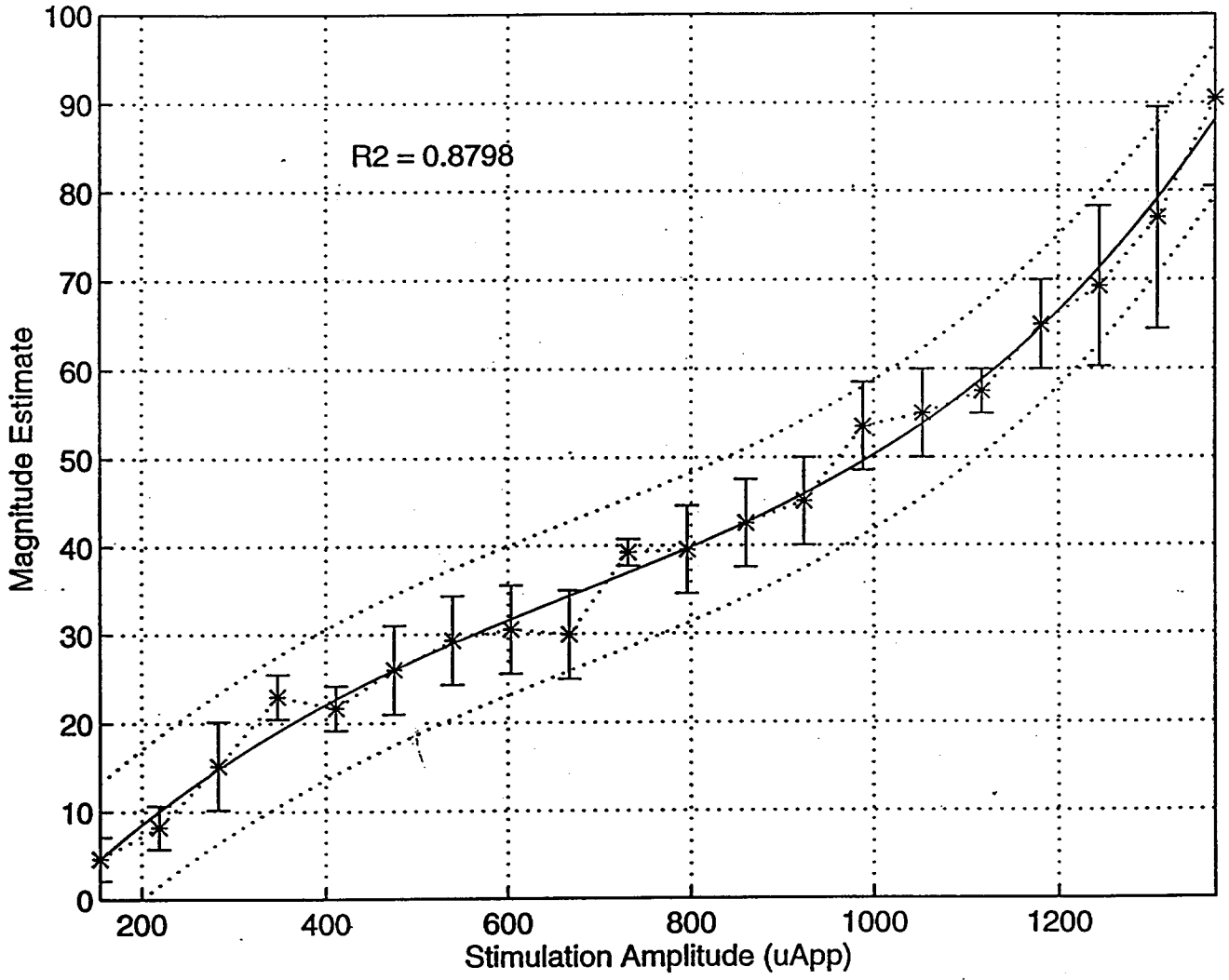


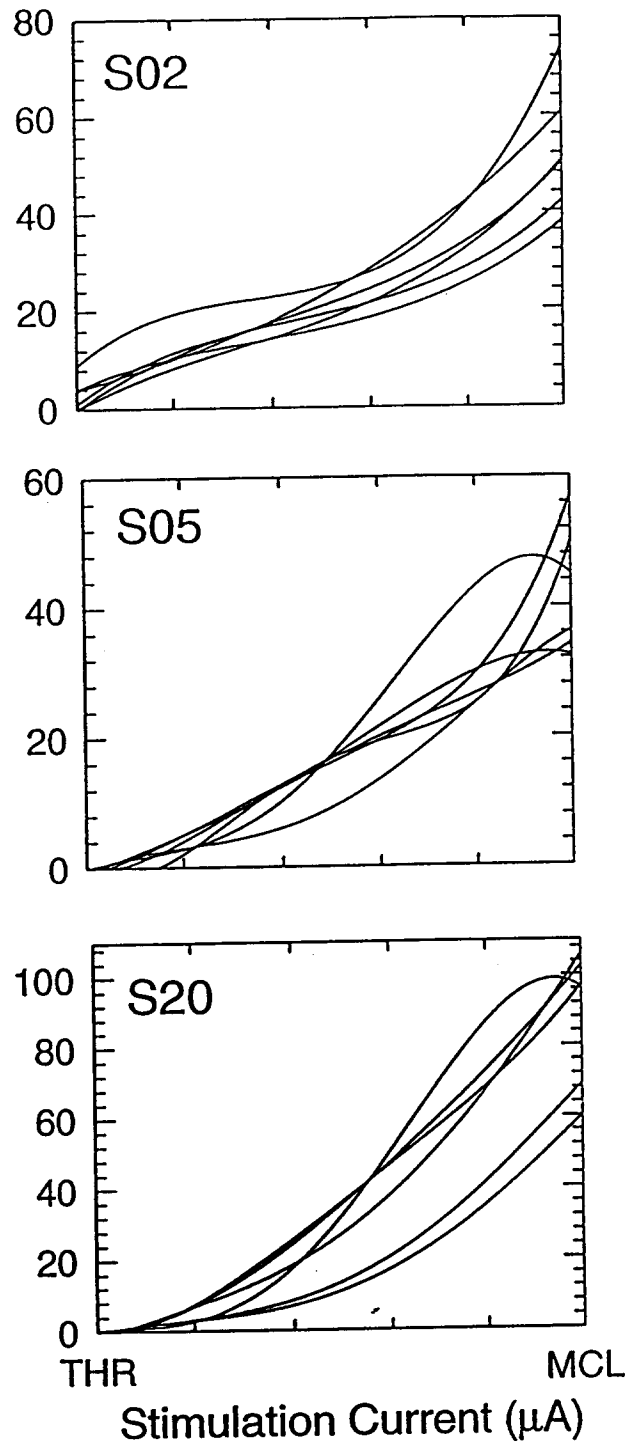
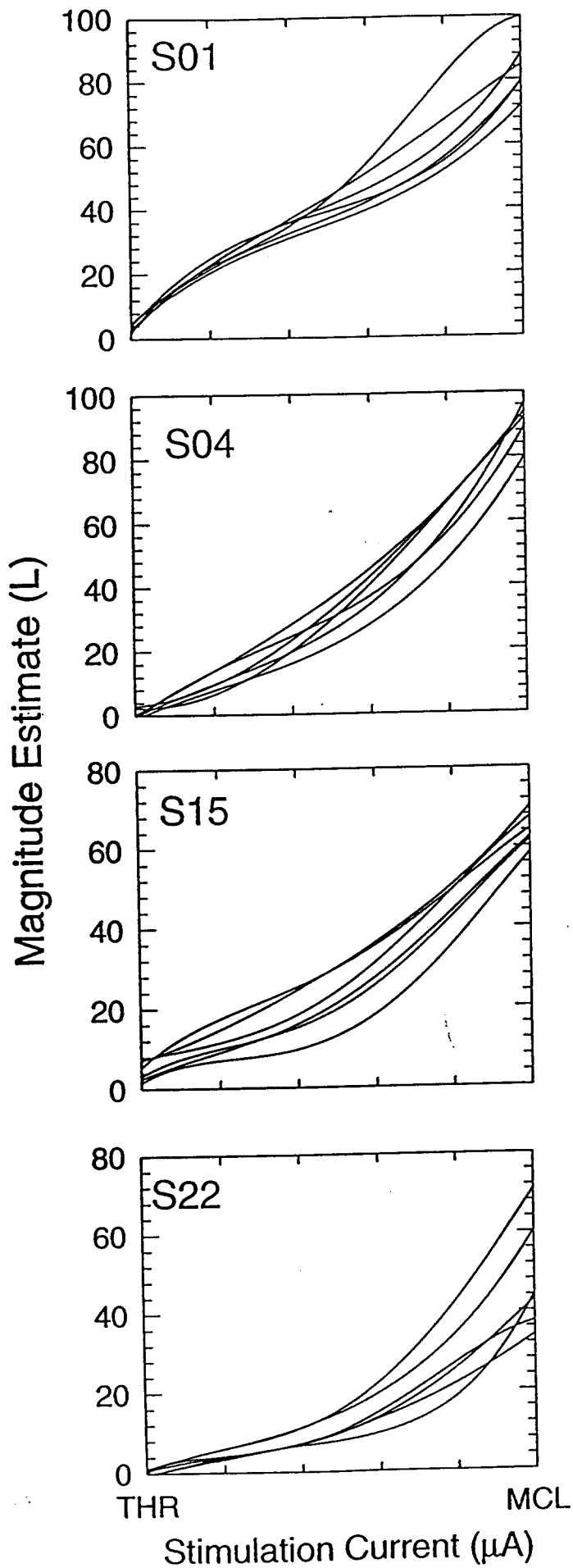
Consonant Identification



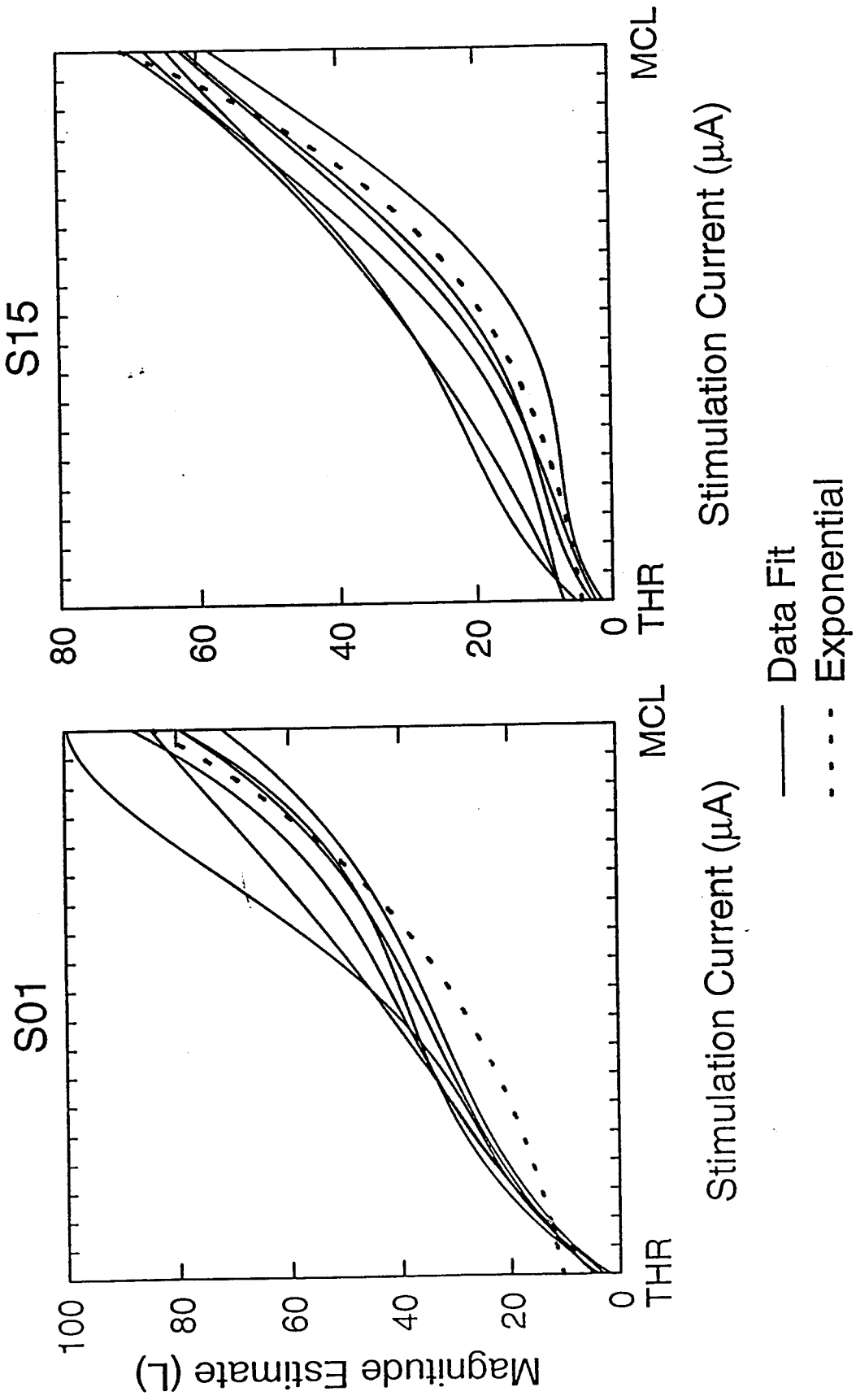
 alpha-Log Map
 L-Log Map

S01 - El.1

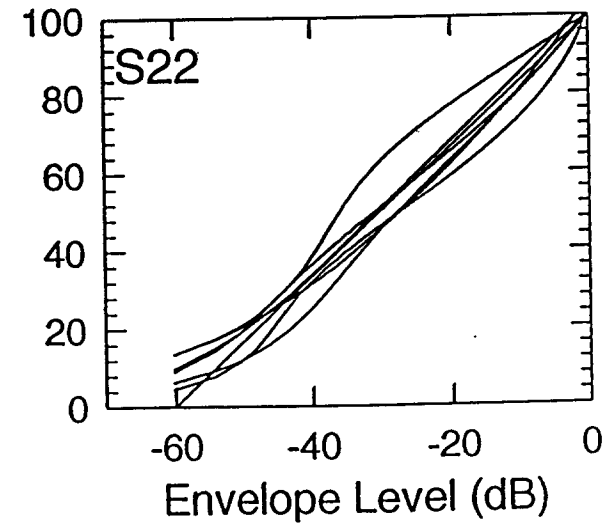
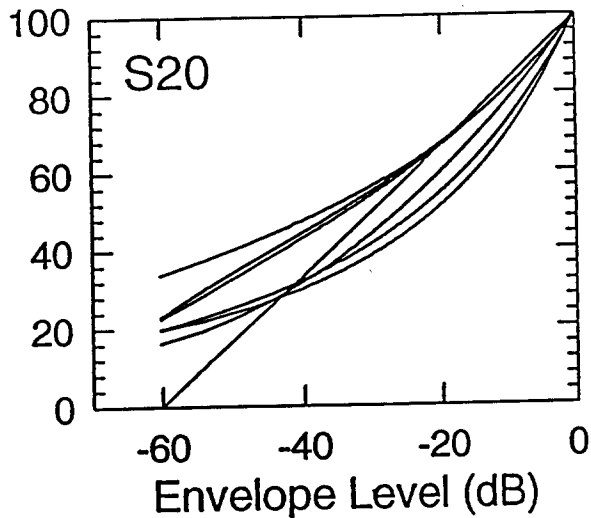
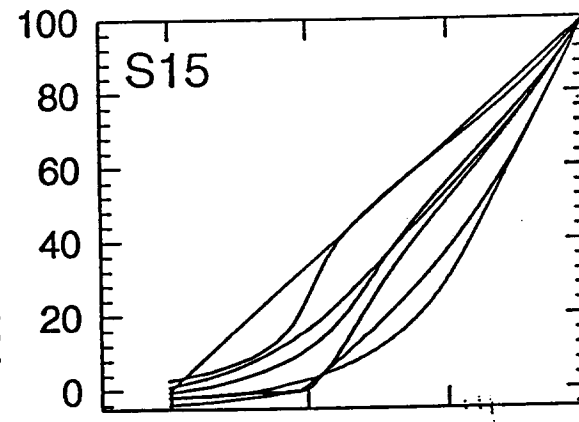
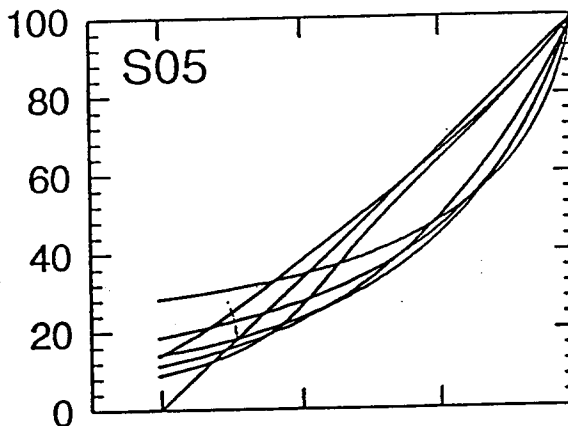
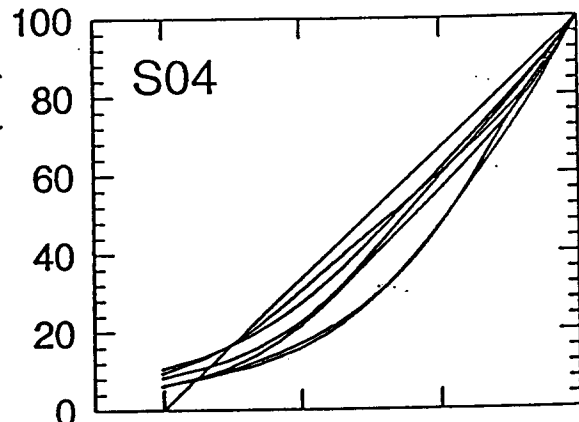
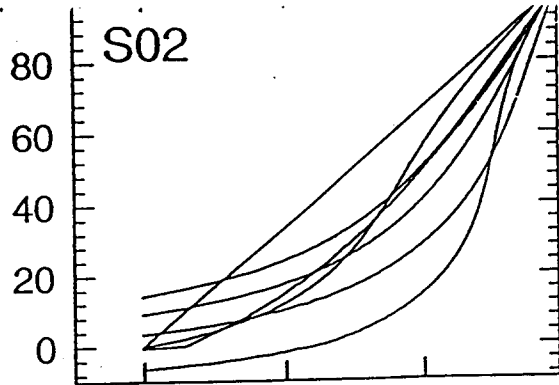
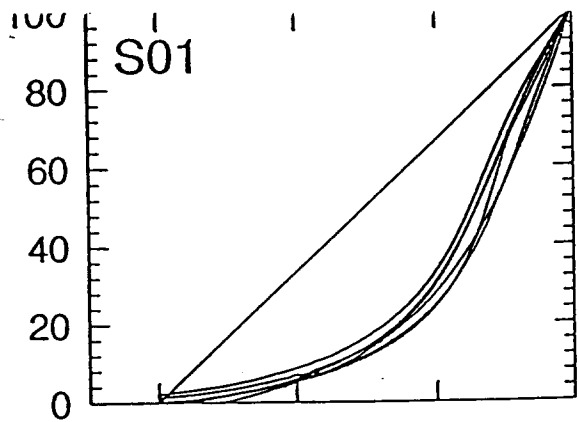


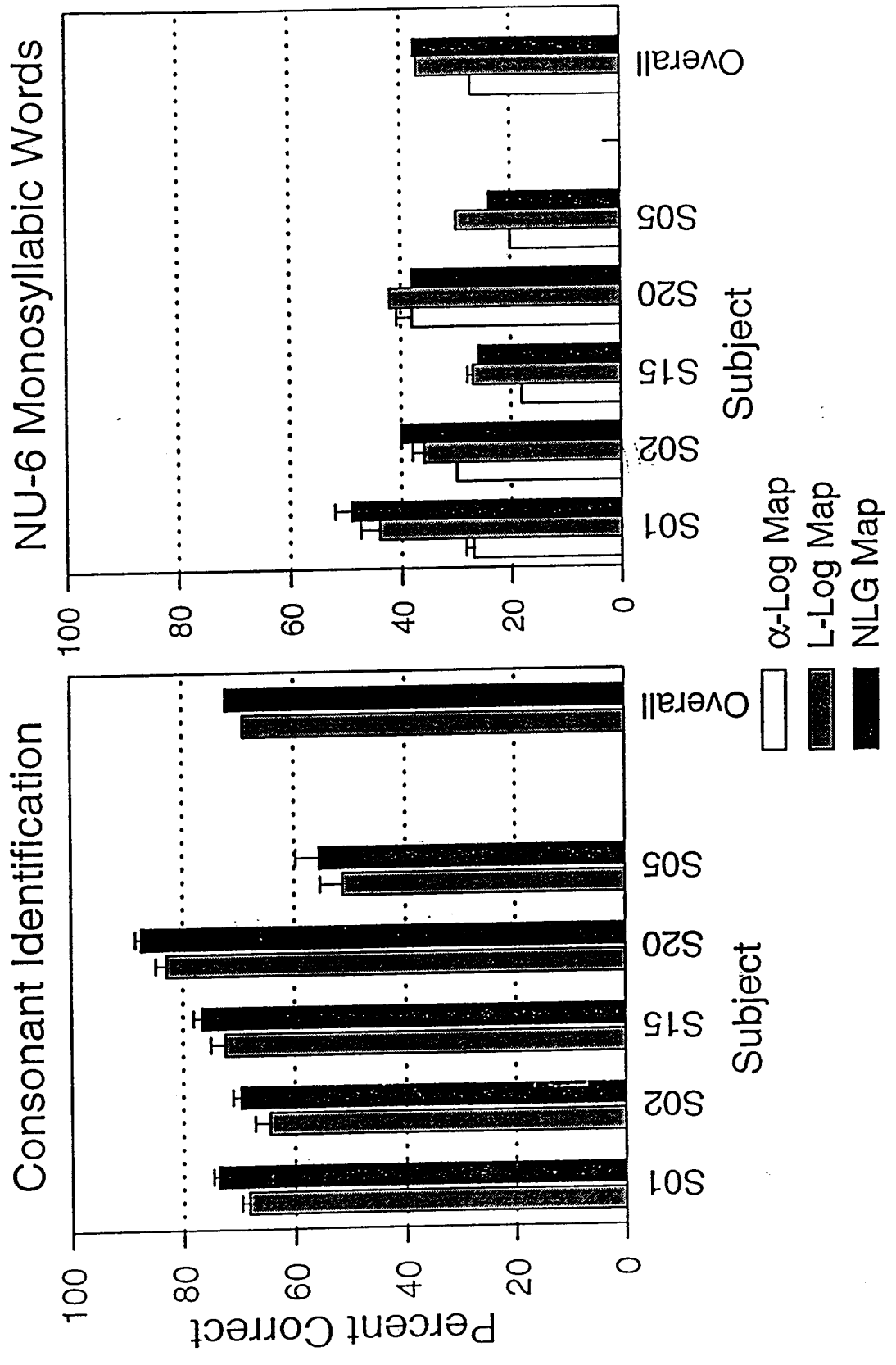


Loudness Growth Function

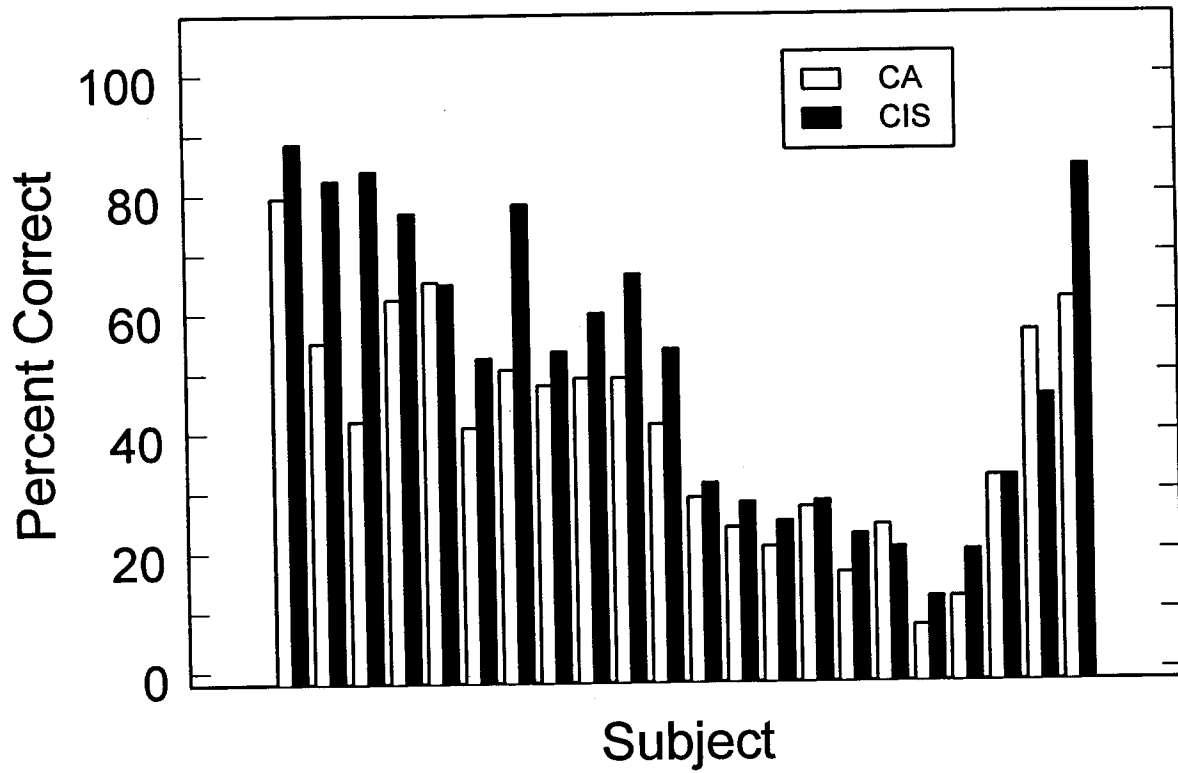


Normalized Stimulation Current (%)

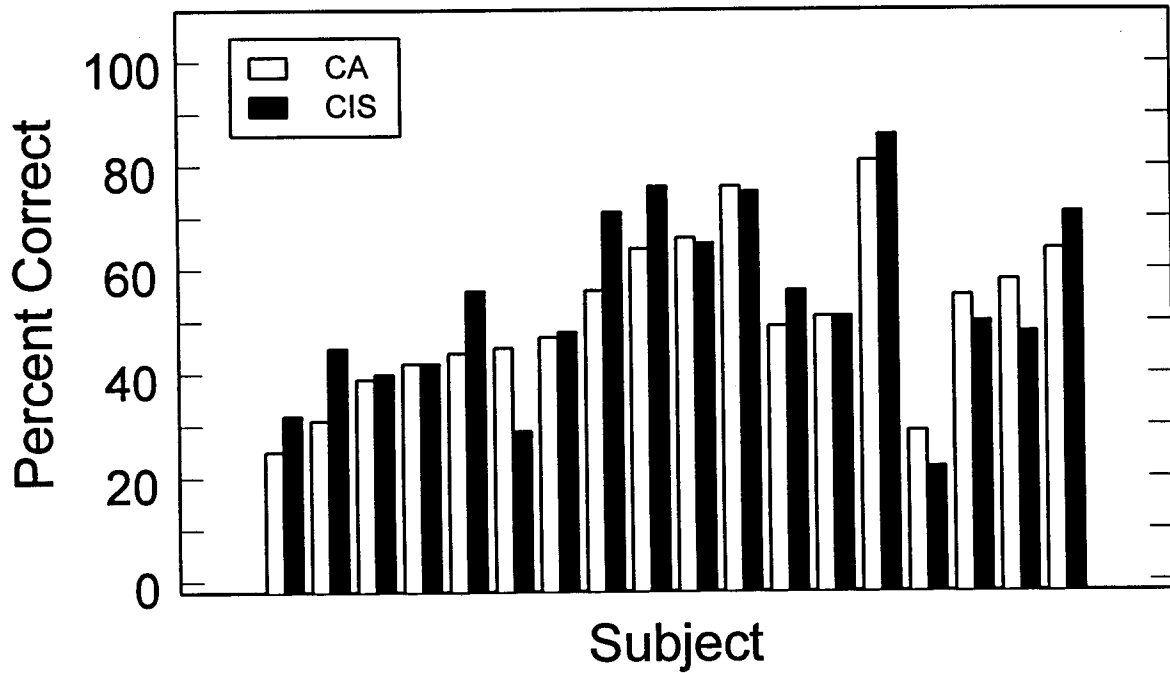




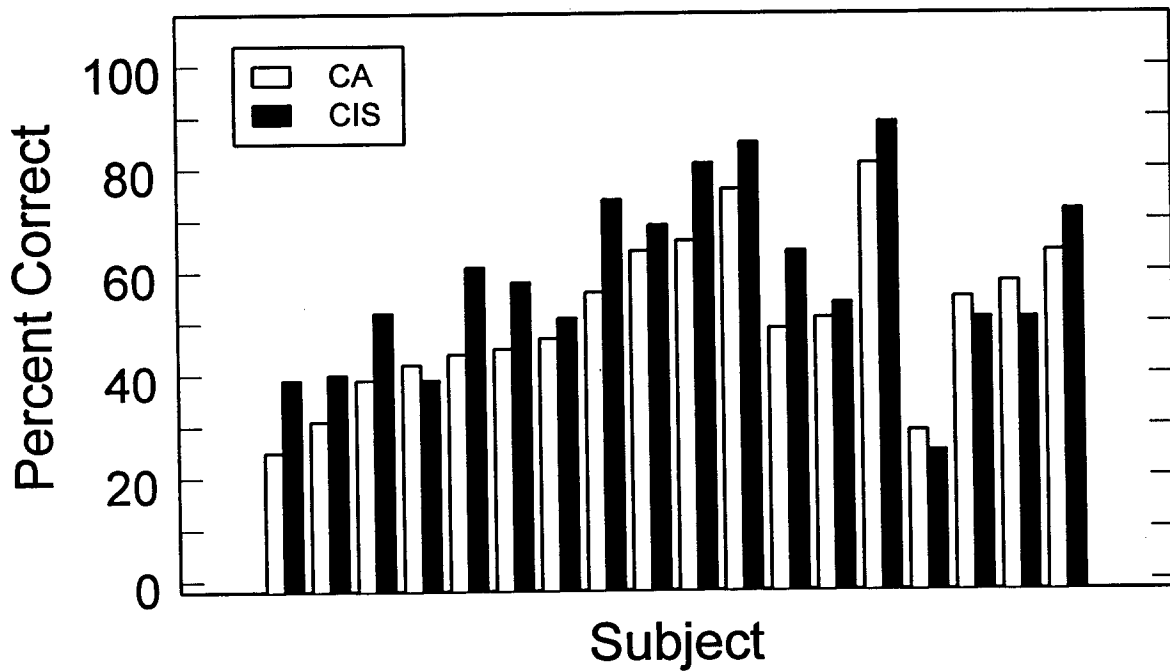
Consonant Recognition (Geneva Subjects)



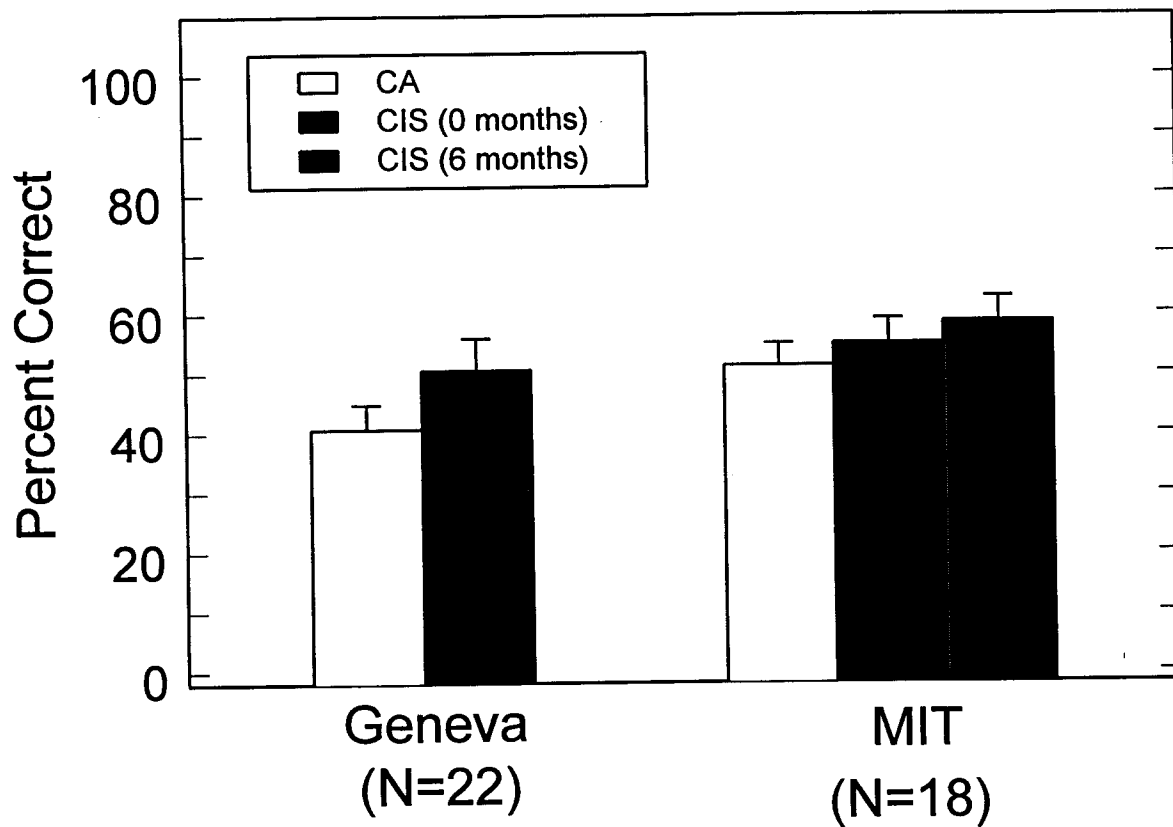
MIT/GWP (0 months)
16 Consonant Test



MIT/GWP (6 months)
16 Consonant Test



Consonant Recognition



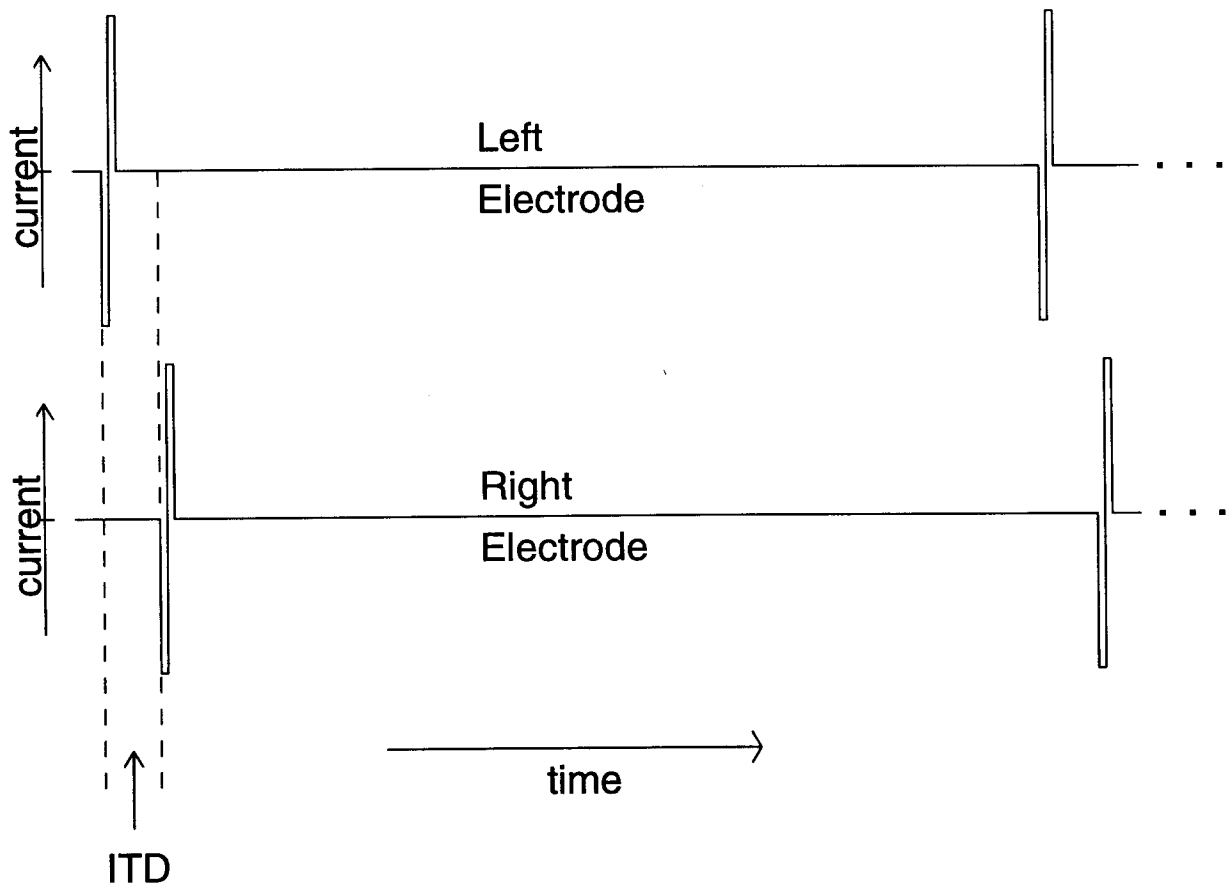


Figure 1: This schematic shows the first two pulse pairs in a bilateral train of pulses. Stimuli delivered to the right electrode (Ineraid) are always delayed relative to those delivered to the left electrode (Clarion). The time axis is to scale and the ITD shown is $600\mu\text{sec}$.

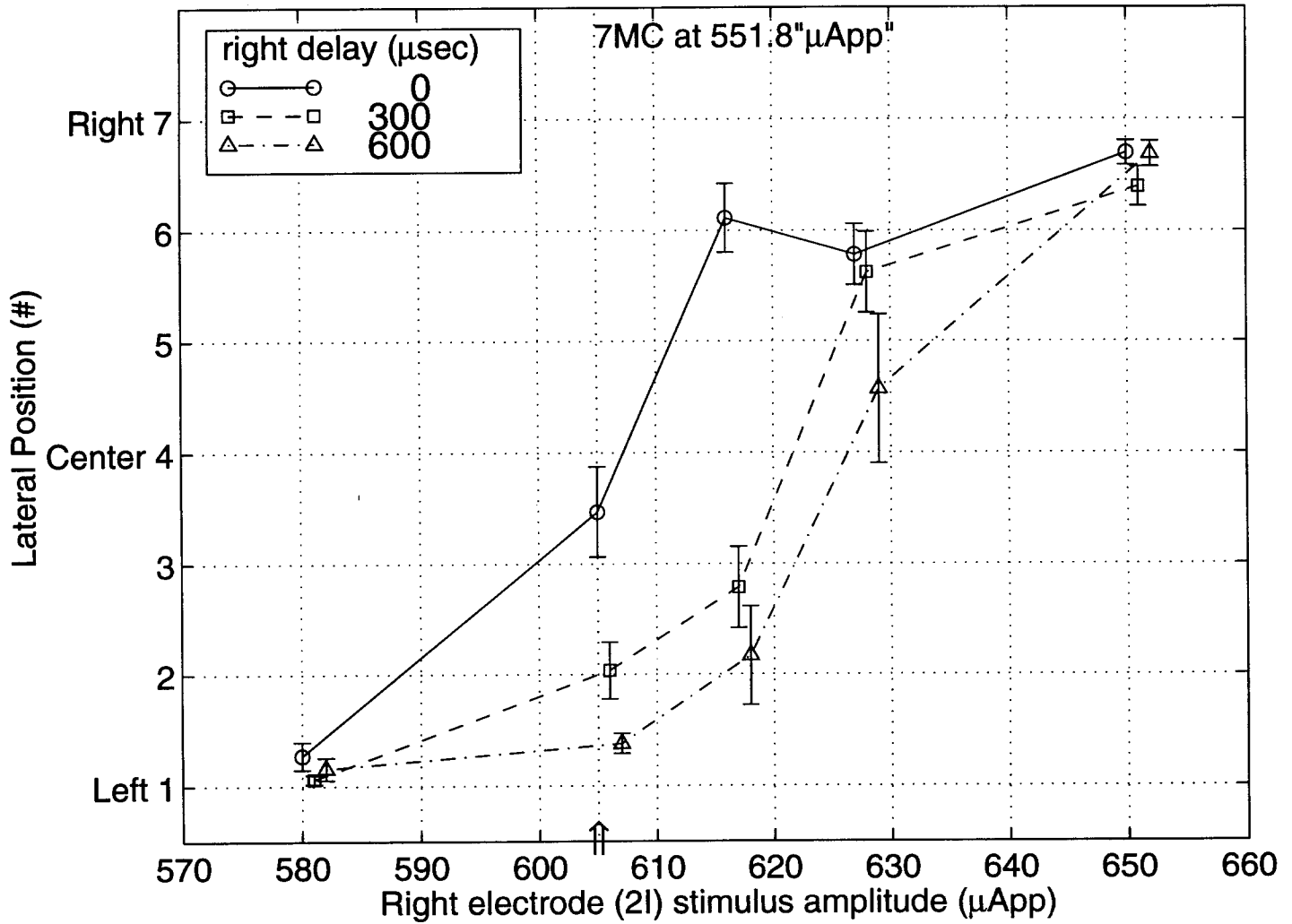


Figure 2: Lateral position as a function of ITD and ILD. This figure shows the mean (\pm standard error) lateral position for the indicated ILD and ITD combinations. Electrode 7MC was held at 551.8 μApp while the amplitude of 2I was varied. The arrow indicates the 2I stimulus amplitude measured during the centering test (see section ??). Non-zero ITD data are shown slightly offset in amplitude for easier visual comparison, although their stimuli were presented at the same amplitudes as the zero ITD stimuli.

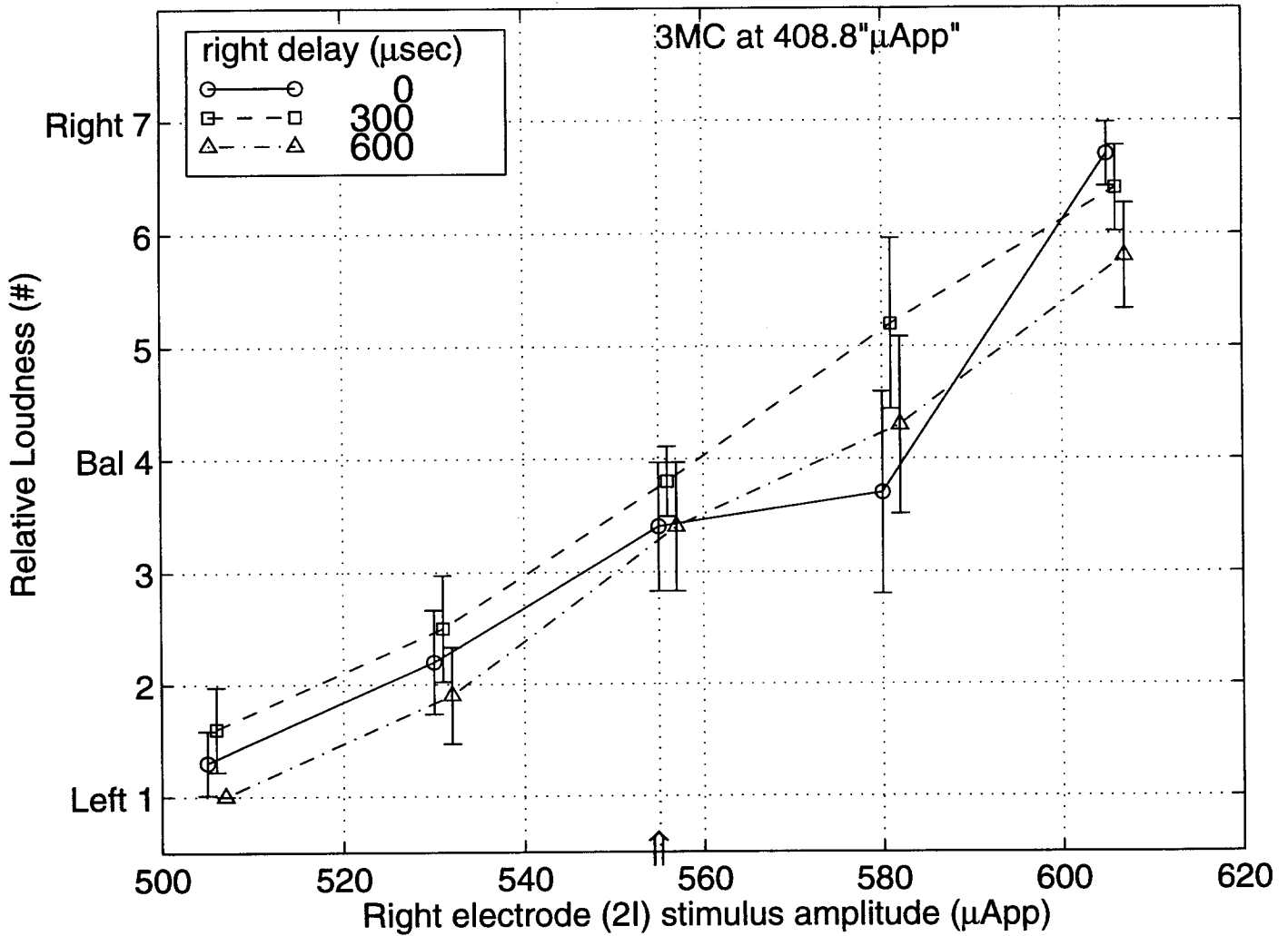
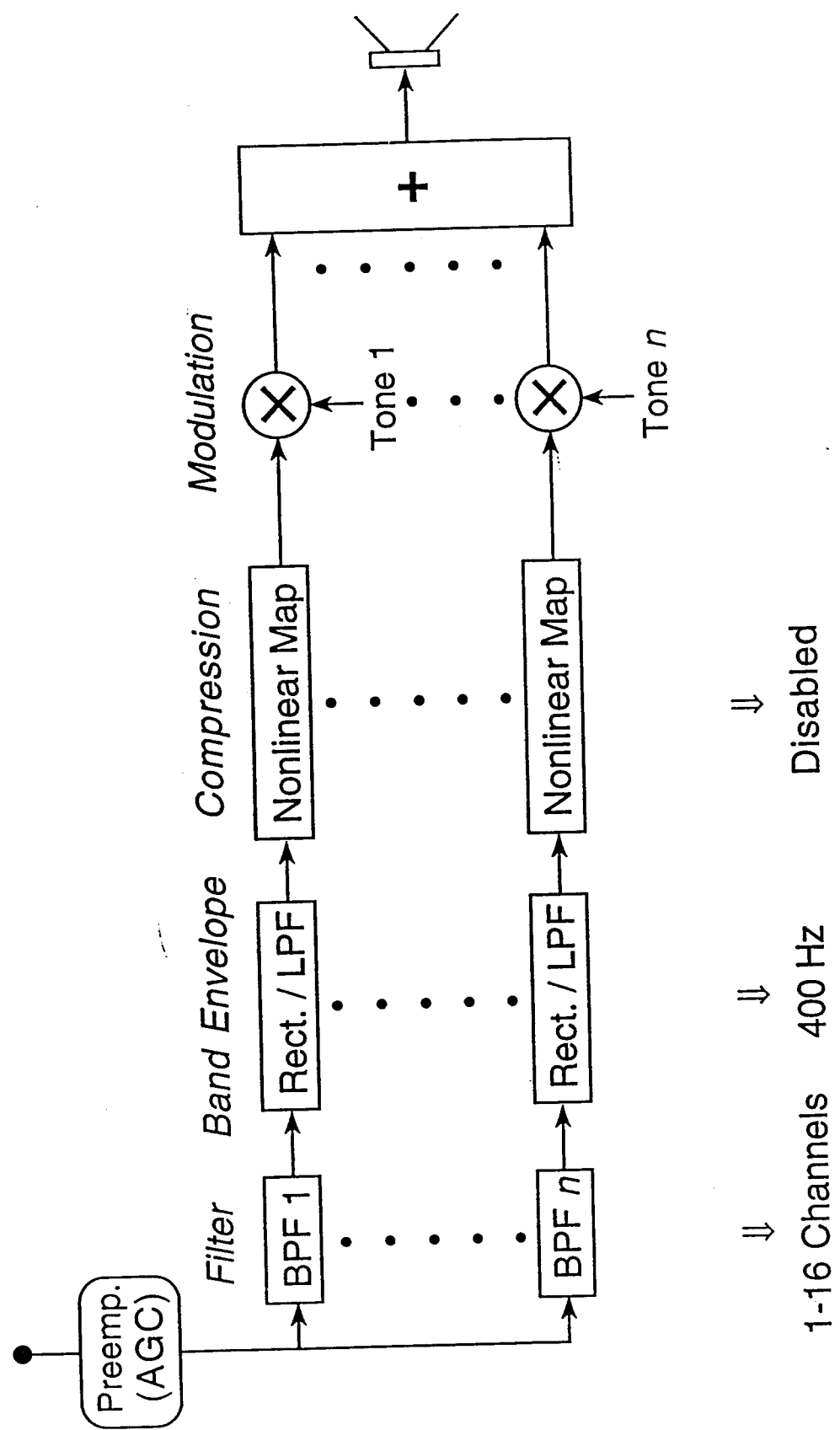
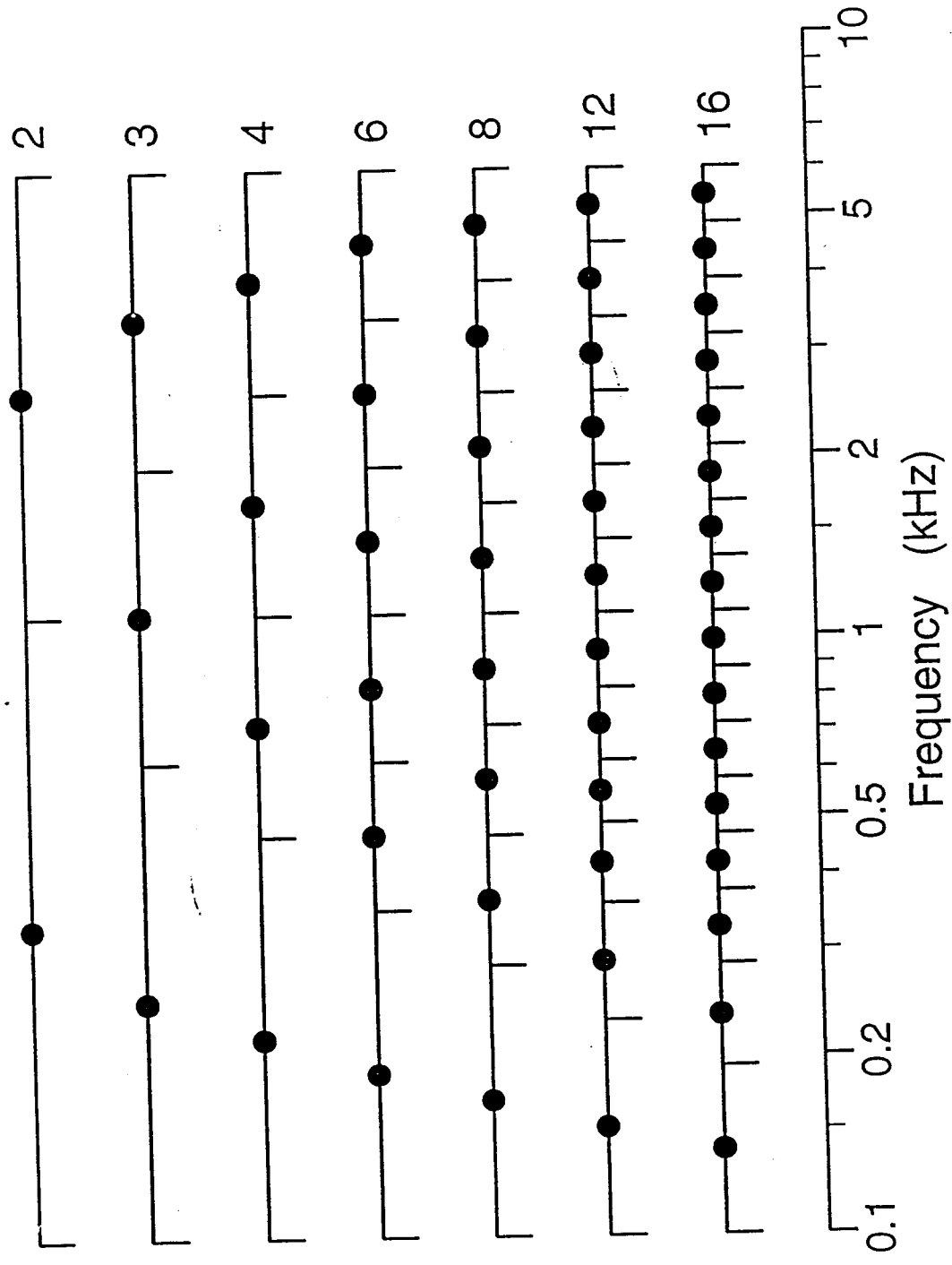


Figure 3: Relative Loudness as a function of ITD and ILD. The figure shows the mean (\pm standard error) relative loudness for 5 ILDs and 3 ITDs. The electrode 3MC amplitude was held at 408.8 "μApp" while the stimulus amplitude of electrode 2I was varied. The arrow indicates the balanced 2I stimulus amplitude measured during the simultaneous loudness balancing test (see section ??). Non-zero ITD data are shown slightly offset in amplitude for easier visual comparison although their stimuli were presented at the same amplitudes as the zero ITD stimuli.

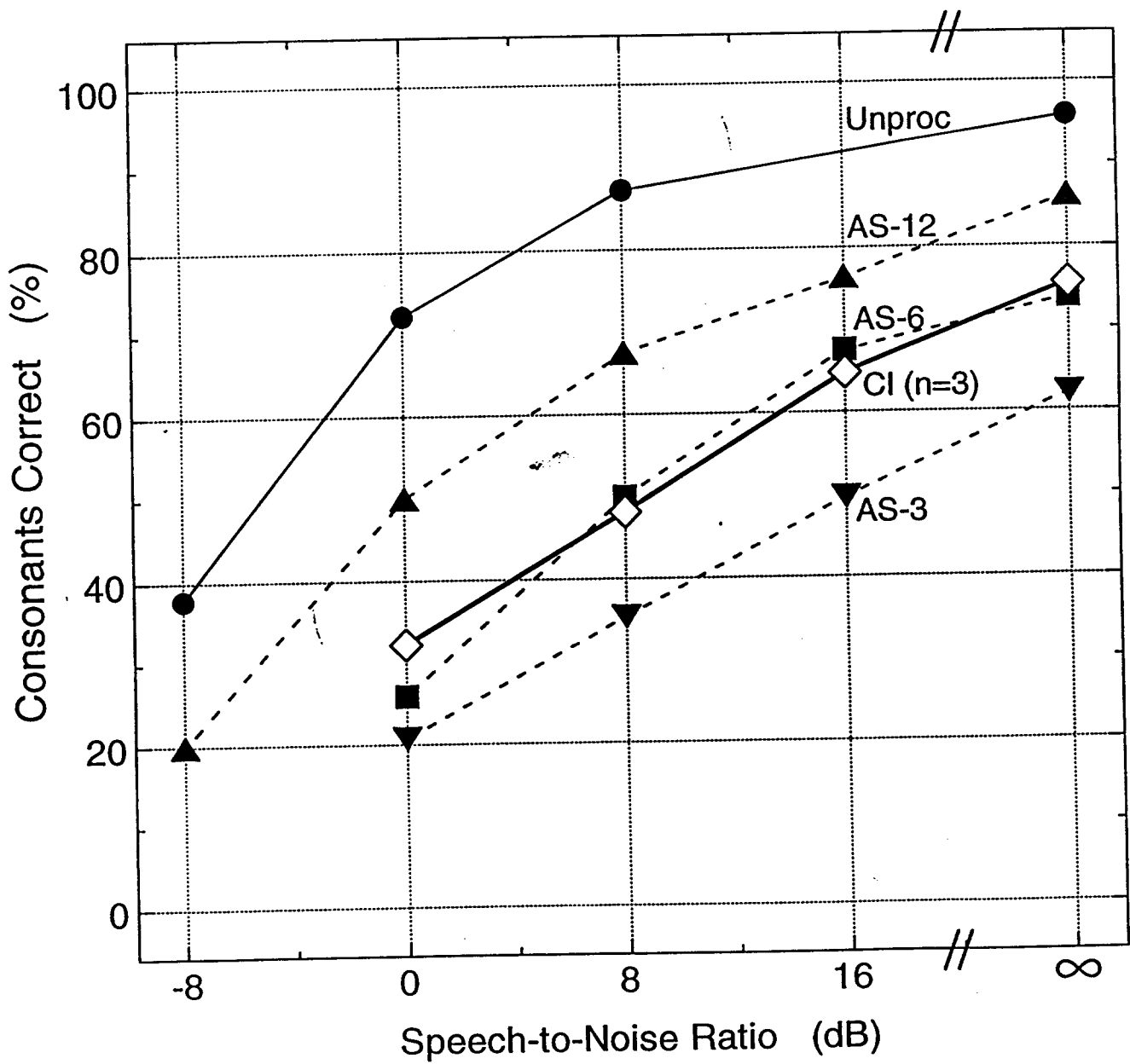
Acoustic Simulations of CIS Processing



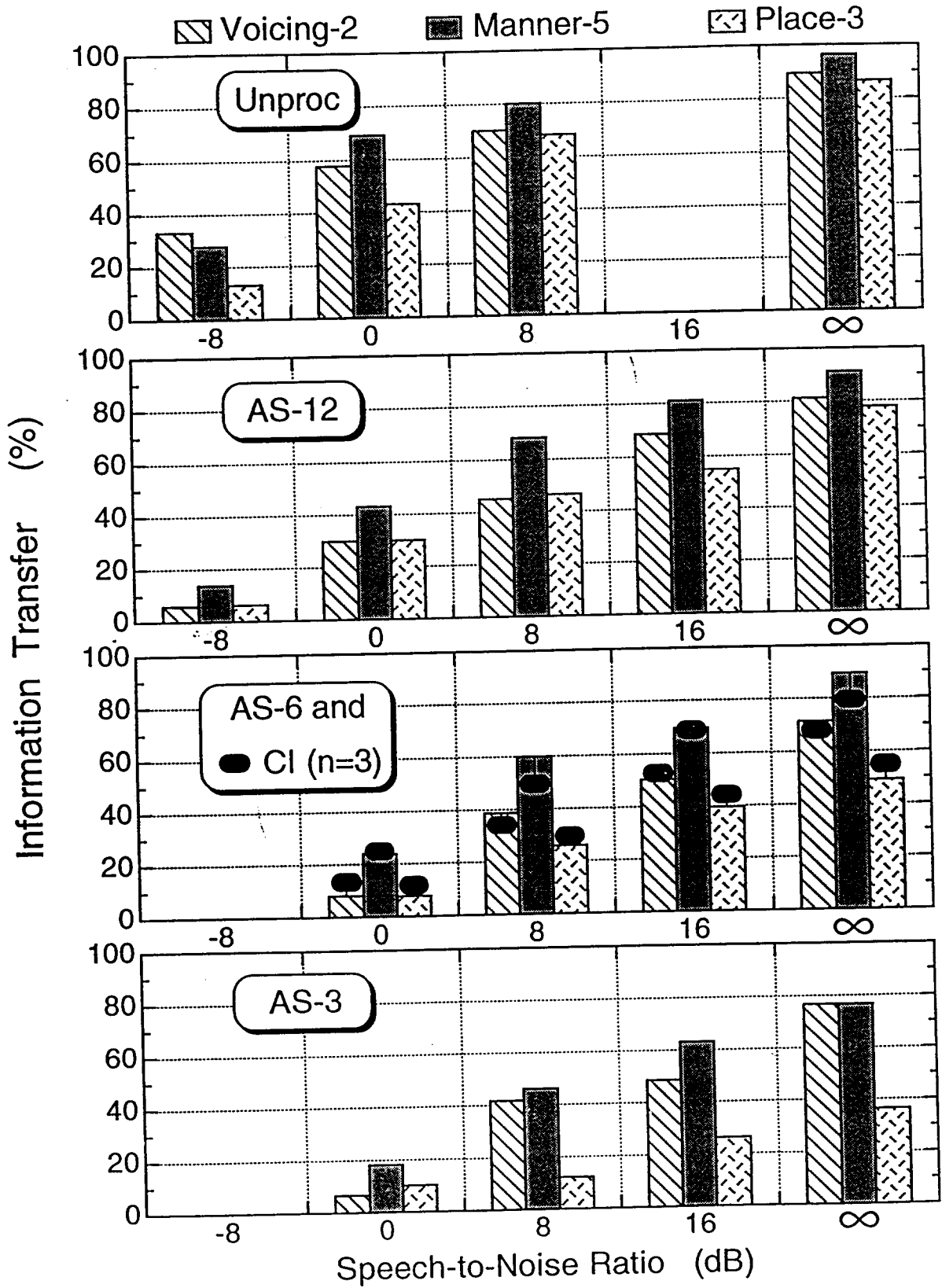
Acoustic Simulation Channels



24 Initial Consonants in Noise

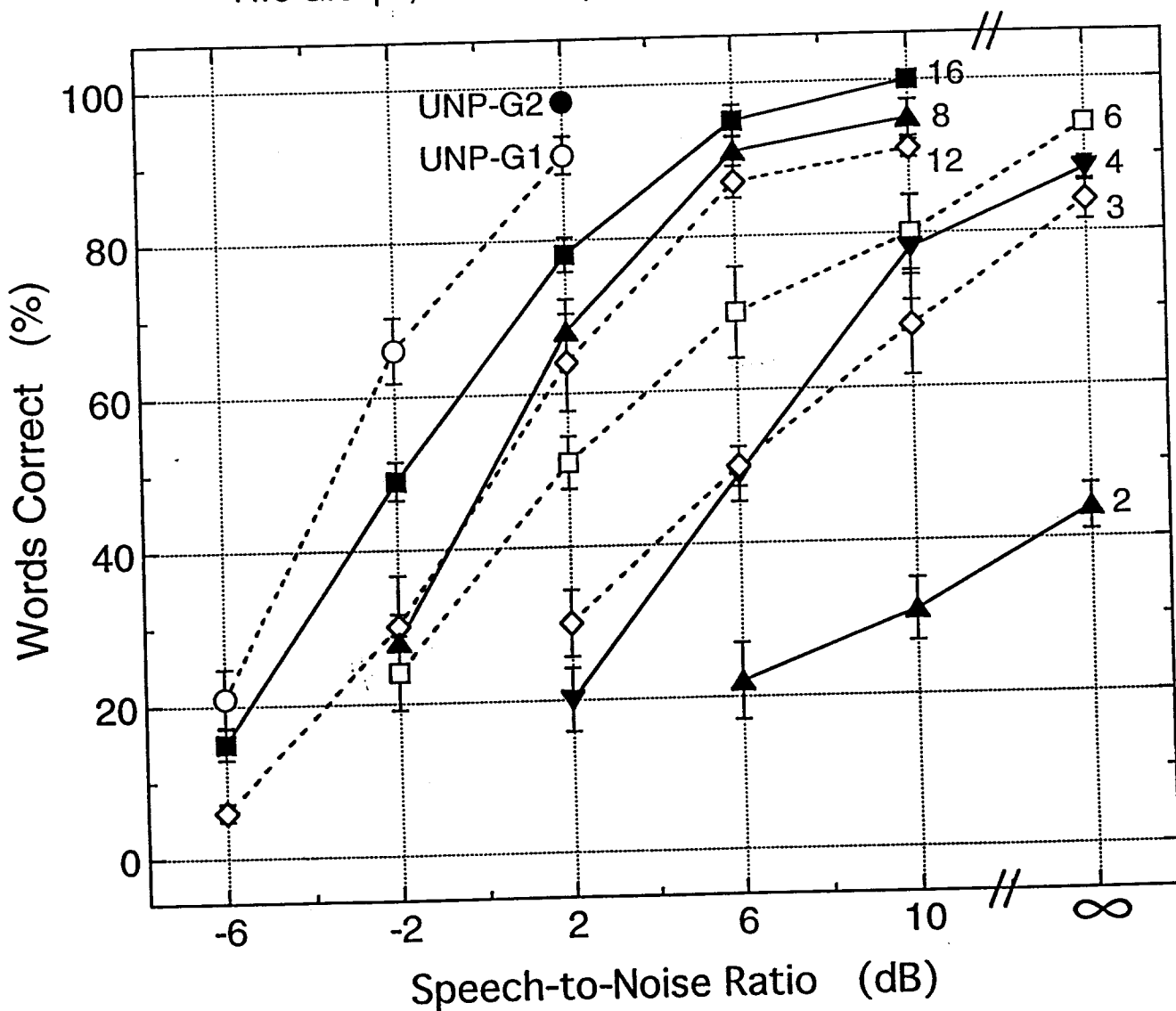


Consonantal Features

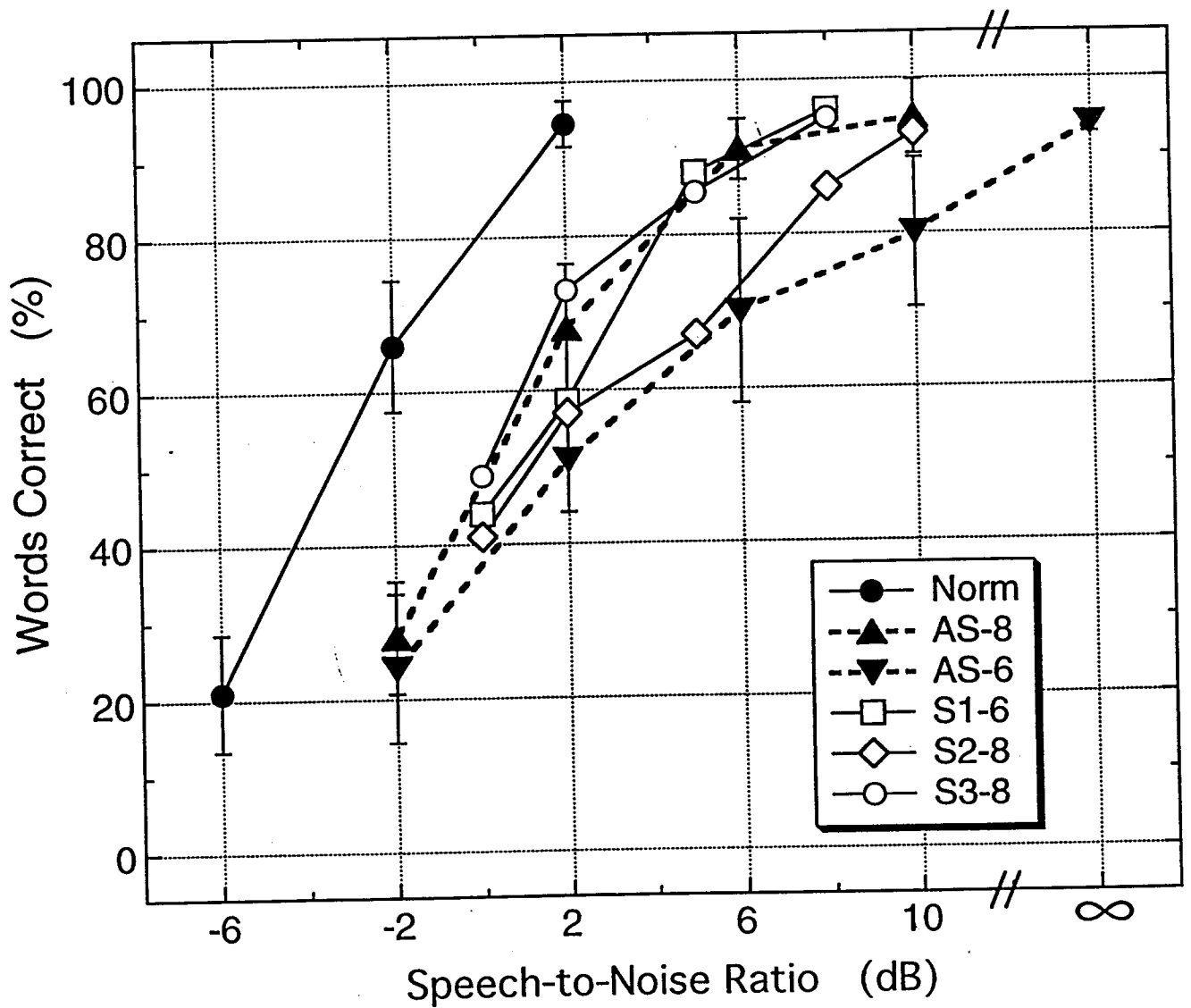


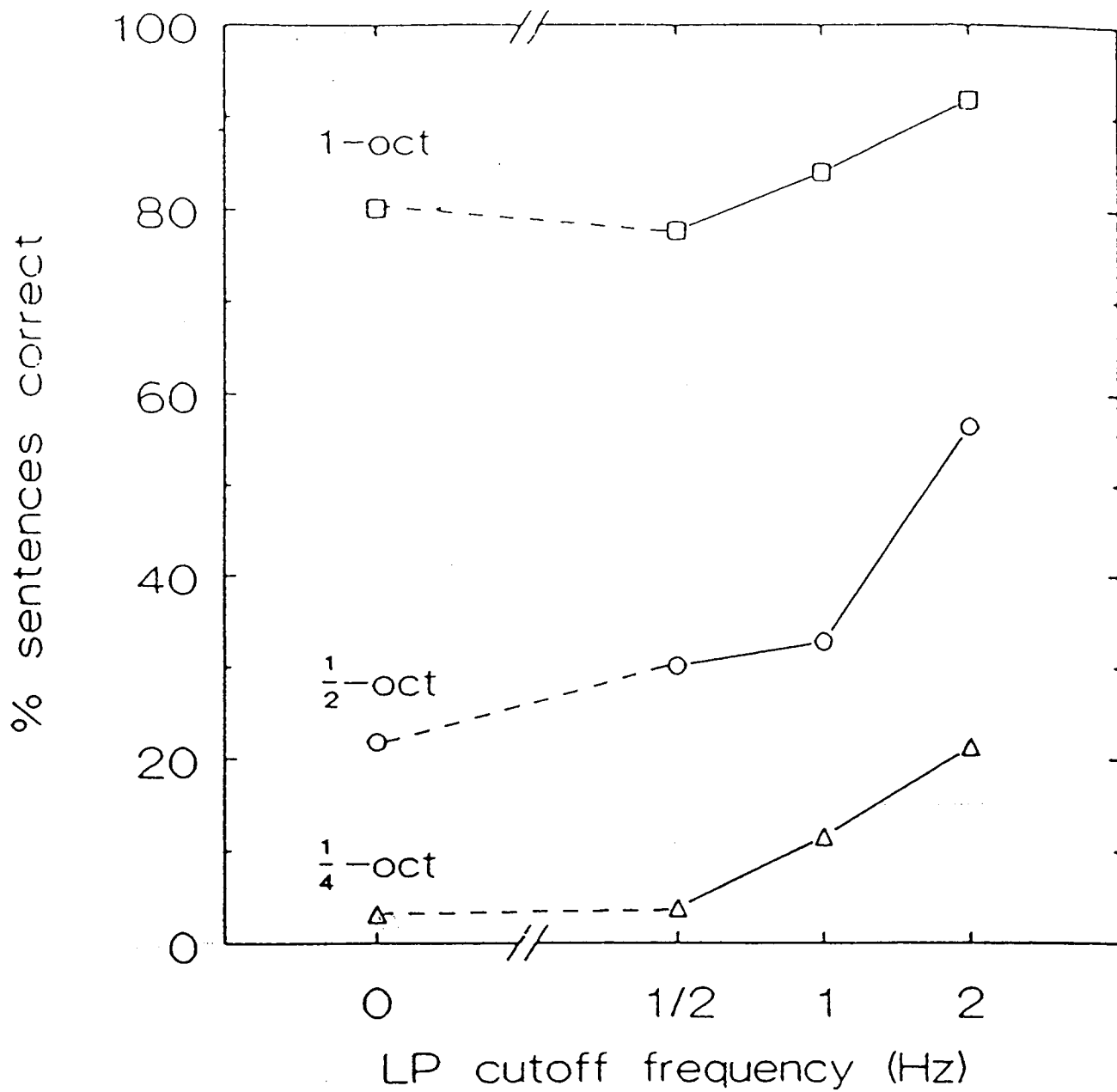
Acoustic Simulations of CIS

- Sentences in Noise (HINT)
- Normal-Hearing Listeners
- Two Groups, $n = 4/\text{Group}$, Mean \pm Std. Error

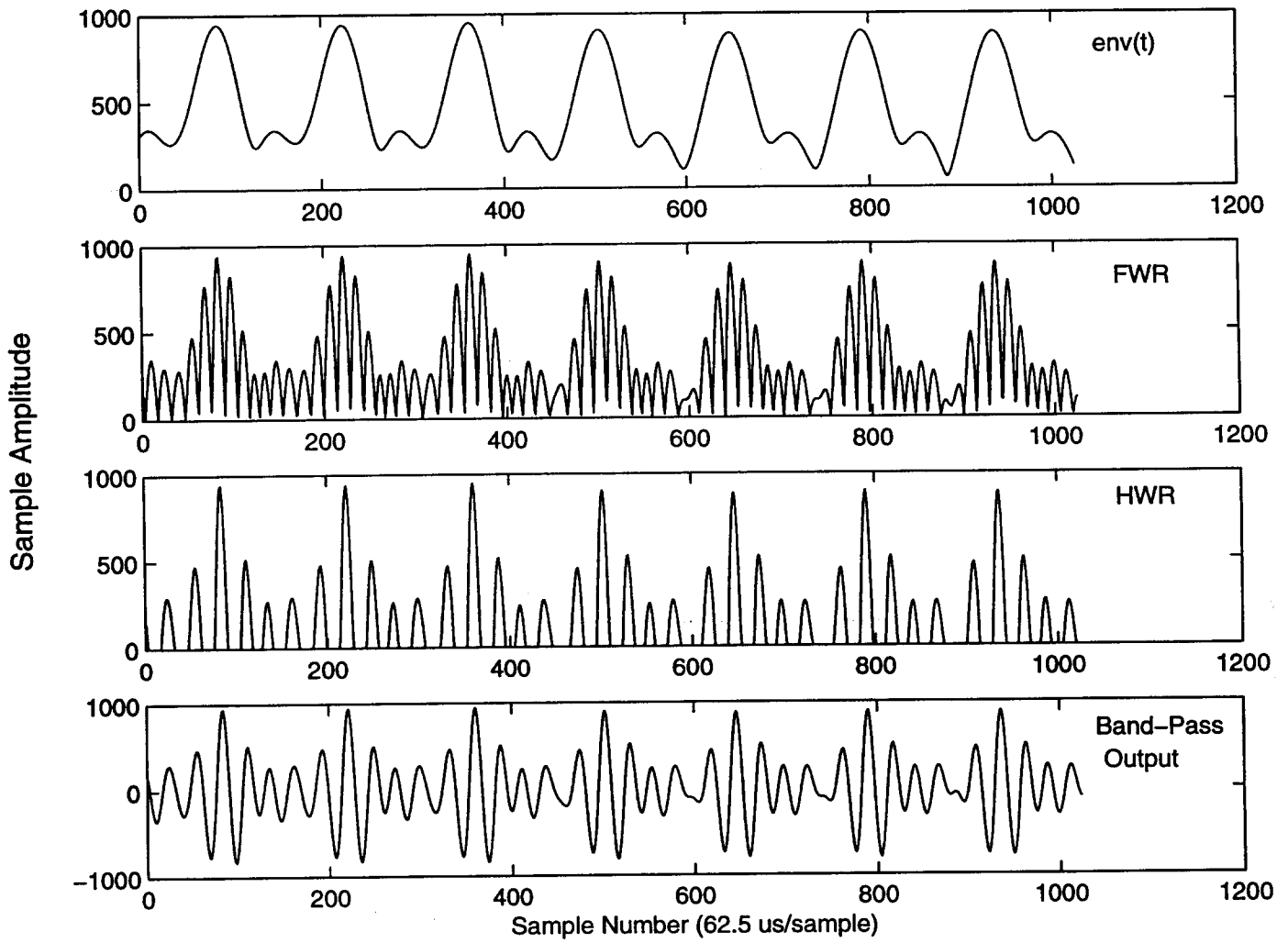


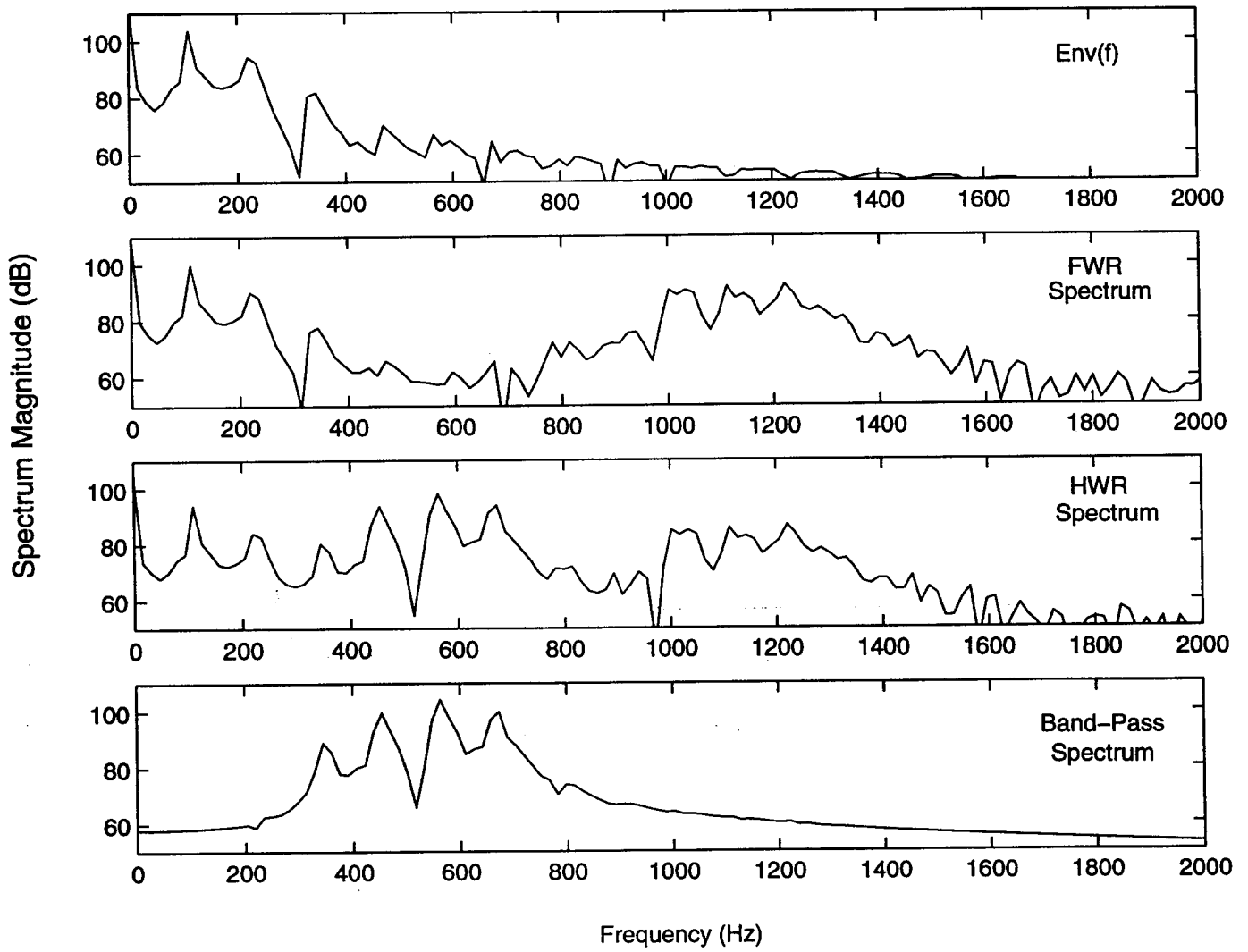
Sentences in Noise (HINT)



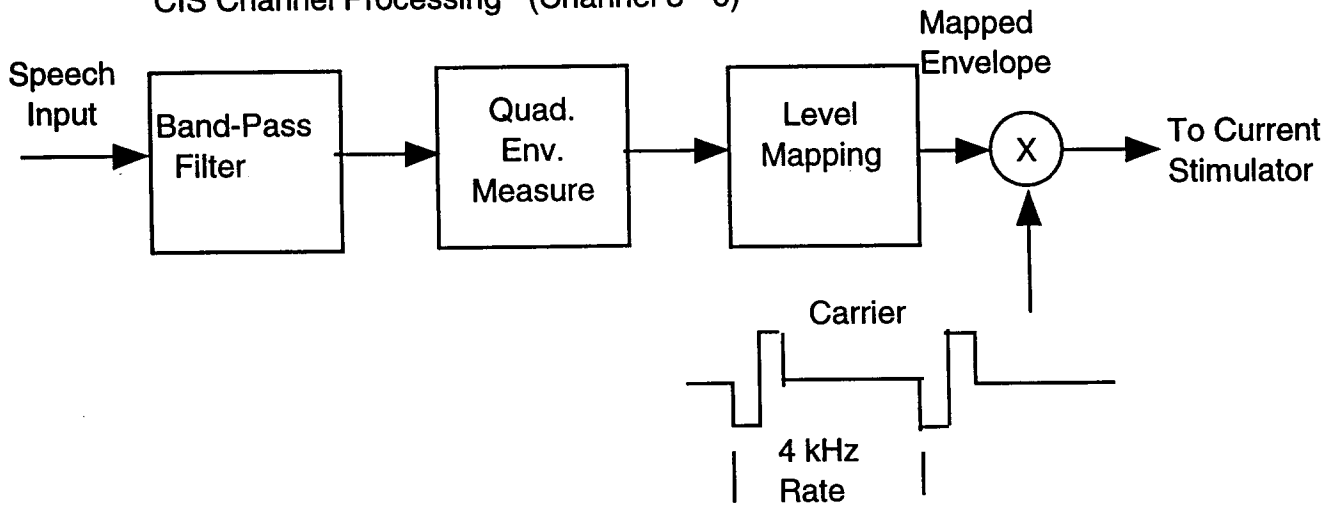


Mean score of sentences in quiet as a function of cutoff frequency, with processing bandwidth as parameter. (AFTER DRULLMAN ET AL 1994)

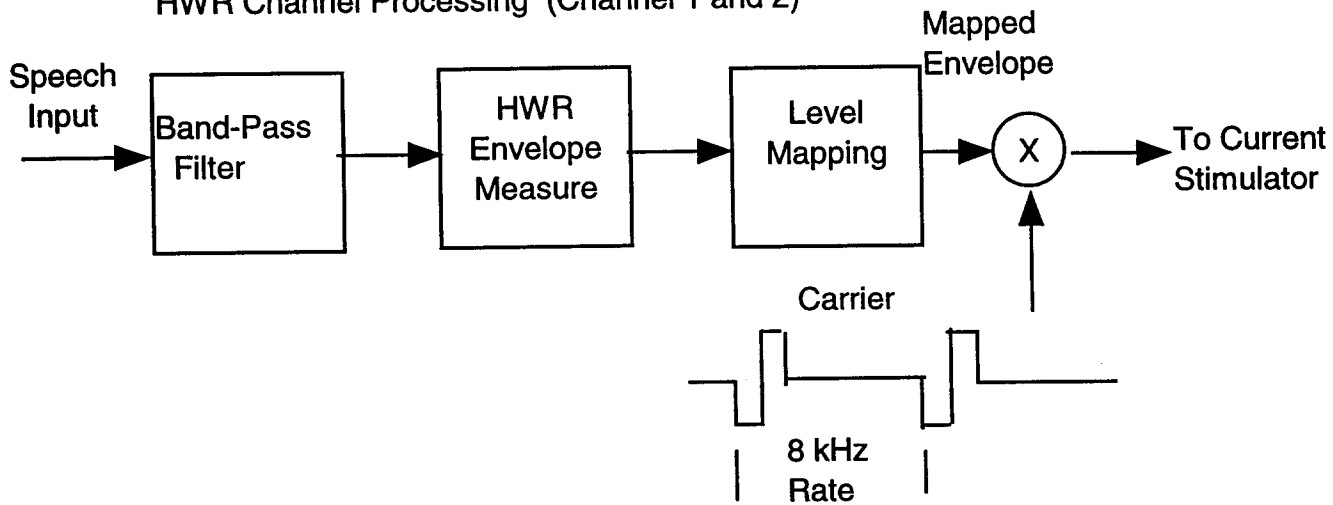




CIS Channel Processing (Channel 3 - 6)



HWR Channel Processing (Channel 1 and 2)



Pulsatile Timing Sequence (Channels 1 and 2 @ 8 kHz, 3-6 @ 4 kHz)

