Third Quarterly Progress Report

July 1, 1996, through September 30, 1996

Speech Processors for Auditory Prostheses

NIH Contract N01-DC-6-2100

submitted by

Donald K. Eddington, PI
Colette Boex-Spano
Lorraine Delhorne
Nico Garcia
Victor A. Noel
William M. Rabinowitz
Joseph Tierney
Margaret E. Whearty

Massachusetts Institute of Technology
Research Laboratory of Electronics
Cambridge, MA

1.0 Introduction

Work performed with the support of this contract is directed at the design, development, and evaluation of speech processors for use with auditory prostheses implanted in deaf humans. Major research efforts are proceeding 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 report concentrates on two of the research areas mentioned above: (1) longitudinal testing using a wearable sound processor developed together with colleagues at the University of Geneva, the Geneva Engineering School and the Research Triangle Institute (RTI) and (2) the refinement of sound processing algorithms for intracochlear stimulation.

2.0 Longitudinal Testing

Last Quarter (Eddington et al., 1996a) we reported preliminary speech reception results from 17 Ineraid subjects who switched from their commercial Ineraid sound processor to the Geneva Wearable Processor (GWP) (Eddington et al., 1996b) running a program developed by us (GWP/MIT) that implements a version of the Continuous Interleaved Sampling (CIS) algorithm (Wilson et al., 1991). Since the last report, an additional Ineraid subject switched to the GWP/MIT system. Some of these subjects have now been wearing the new system for more than six months.

Figure 1 shows single-syllable word (Panel A) and consonant (Panel B) identification scores for 18 subjects using their commercial Ineraid sound processor and their GWP/MIT system. The results are arranged in ascending order of Ineraid score within each of three subject groups: those subjects tending to score higher with the interleaved processor, those scoring about the same with the two processors and those scoring higher with the Ineraid. Moving from left to right, the GWP/MIT bars for each subject represent scores obtained at time 0, 1, 3 and 6 months from the day they switched from the Ineraid to the GWP/MIT system.

NU6 word scores tend to initially increase with time for most of the subjects. Notice that 8 of the 13 subjects showing higher NU6 scores for the GWP/MIT system initially scored worse than with their Ineraid processor. In the case of consonant identification, trends of performance improving with time are not as obvious. In fact, the number of subjects showing an upward trend with time in consonant identification is about the same as the number showing a downward trend. Because of the preliminary nature of these results, we have not completed analyses to determine whether any of these "trends" are significant. That will be reported in a subsequent QPR when we review the full battery of longitudinal tests conducted with these subjects.

In Figure 2, the average word (panel A) and consonant scores (panel B) obtained with the GWP/MIT system are plotted as a function of the Ineraid score for each subject. Application of the ANOVA model to these results indicates that NU6 word identification is significantly better (0.05 level) for the GWP/MIT processor. The consonant data do not show a significant processor effect.

While only 7 of the 18 subjects demonstrate statistically significant (0.05 level), individual improvements in either consonant or NU6 word recognition, 14 report a strong, subjective preference for the GWP/MIT sound processor. These individuals now consider their Ineraid sound processor to be a very inferior back-up system. Of the four subjects not showing a strong preference for the GWP/MIT system, only S15 reports a significant drop in both sound quality and intelligibility that is consistent with her drop in NU6 and consonant scores. Subjects S05, S09 and S21 do not notice a significant difference in overall intelligibility between the GWP/MIT and their Ineraid system.

3.0 Level Mapping

Because the dynamic range of sound is at least four orders of magnitude larger than the dynamic range of electrode stimulus level for most cochlear implant users, the method used to map acoustic level to electric stimulus level is a key component of any sound processor. In previous QPRs (Eddington et al., 1996a; Eddington et al., 1996b), we described the mapping strategy of the standard interleaved processor used in both our laboratory and GWP studies. For convenience, this mapping strategy is summarized here.

3.1 Standard MIT Mapping

Figure 3 shows the block diagram of the mapping stage located between the envelope detector and the voltage-to-current converter of each CIS channel. The input gain (Gin) determines the range of envelope levels that will fall within the 60 dB input range that is mapped. Gin 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) database of speech materials played at a conversational level.

Initially, the output dynamic range (DR in Figure 3) is determined for each channel based on the psychophysical threshold (THR) and most comfortable (MCL) stimulus levels measured using the electrode associated with that channel. The stimulus waveform used to make these measurements roughly corresponds to that used by the processor of interest (e.g., biphasic, cathodic-first pulse train, 2000 Hz repetition rate, 31.25 μs/phase, 50 ms duration in our current five and six-channel CIS implementations). For a given Gin, the DR is adjusted to produce a comfortable listening level by varying a scale factor, SF, where DR = (SF*(MCL-THR)+THR)/THR.

The shape of the function that maps the input range to the output range is logarithmic in our current laboratory and GWP/MIT sound processing systems. This relationship was an initial selection based on the work at RTI where a wide range of mapping functions were explored. Figure 4 compares magnitude estimates made by an implantee using the standard logarithmic mapping function with a line representing magnitude estimates of normal hearing listeners for a 1 kHz tone. (The methods used to obtain the magnitude estimates are given in the figure caption.) Results like these indicate that loudness grows very differently for implant subjects using logarithmic mapping functions than it does for normal hearing listeners. In order to determine whether sound processing systems that more closely approximate normal loudness growth will improve speech reception for implantees, we are developing methods to guide the design of mapping functions for CIS processors that restore the normal growth of loudness.

3.2 Mapping Functions that Restore Normal Growth of Loudness

A relatively straight-forward method for designing "normal loudness growth" (NLG) mapping functions is based on magnitude estimates of normal subjects in response to single-tone stimuli. This method uses the relationship between loudness and the sound pressure of single tones measured in normal hearing subjects and the relationship between loudness and electric stimulus level measured at individual electrodes in implantees to derive a set of mapping functions for a CIS processor that restores normal loudness growth for single tones at the processor channels' center frequencies.

The first step of this process is to obtain loudness growth functions for each of the implantee's electrodes. Figure 5 shows an example of measurements made to determine this function for Subject S01's electrode 3. The magnitude estimation task is like that described in the caption of Figure 4 and the stimulus waveform was the same described above for the measurement of THR and MCL except that the duration is 300 instead of 50 ms. Once the magnitude estimates are obtained for each electrode, they can be fit to give a loudness growth function, $L_e(I)$, of the form

$$L_e = a_{0,e} + a_{1,e}I + a_{2,e}I^2 + a_{3,e}I^3 + a_{4,e}I^4 \qquad \text{Eq. 1}$$

where L_e is the loudness produced by stimulating electrode "e" at stimulus level I. A maximum current is selected for each electrode ($I_{e,max}$) that corresponds to a loudness magnitude (L_{max}) that is constant across electrodes. Next, using the normal loudness growth function $L=kP^{0.6}$, the relation between the logarithm of the ratios of loudness and sound pressure (P) can be expressed as

$$\log(L_i/L_{max}) = 0.6 * \log(P_i/P_{max}) \qquad \text{Eq. 2}$$

Because the sound pressure (P) of a tone at the center frequency of channel "c" is linearly related to the envelope amplitude in the channel (E_c), Eq. 2 can be rewritten as

$$L_{c,i} = L_{max}*(E_{c,i}/E_{max})^{0.6} \qquad \text{Eq. 3}$$

Using Eq. 3 and Eq. 1, a mapping function can be computed that specifies the relationship between current level and envelope amplitude, $I_c(E_c)$, that produces normal loudness growth for a tone presented at the center frequency of a particular channel. The resulting mapping function for S01's channel 3 (connected to electrode 3) is shown together with the standard logarithmic mapping function in the top panel of Figure 6. It is clear that the difference between the two is very large.

In order to test the method described above, we measured magnitude estimates for tone stimuli in subjects using CIS processors with NLG mapping functions and compared them to normal loudness growth functions. The bottom panel of Figure 6 shows a representative example of such results. Magnitude estimates obtained from subject S01 using a 1300 Hz tone burst (center frequency for channel 3) are plotted for both the standard mapping function (open symbols) and the mapping function designed to restore normal loudness growth (closed symbols). Notice that the overall growth of loudness exhibited by the magnitude estimates made using the NLG mapping function is very similar to the growth of loudness for normal listeners represented by the dashed line.

While the NLG mapping functions derived using this method are based on single tones at a channel's center frequency, we thought it would be interesting to test their effect on speech reception. The table below shows preliminary identification results from two subjects tested with the Iowa, 16-Consonant Test (Tyler et al., 1987). The mean scores shown represent at least 15 randomized presentations of the 16 consonant lists. Subject S01's scores are about the same across conditions for these acute, laboratory tests. Although the scores for Subject S22 show a marked improvement for the NLG mapping, this result must be considered very preliminary since the THR and MCL parameters used to define the standard mapping function were measured several months earlier.

16-Consonant Identification		
Subject	Standard Logarithmic Mapping (Percent Correct)	Normal Loudness Growth Mapping (Percent Correct)
S01	77	80
S22	65	81

4.0 Future Work

Next Quarter we plan to continue work directed at NLG mapping functions. We will begin adding a second software program to the Geneva Wearable Processor (GWP) of five to seven subjects that includes a NLG mapping function. We also plan to conduct additional laboratory studies using NLG mapping functions to determine how loudness grows for tone frequencies positioned at the boundary of two processor channels and for broader band inputs like speech and speech-shaped noise. We hope that these investigations will lead us to refined NLG mapping functions that will improve the listening experience of a large number of subjects.

We also plan to analyze the speech reception results we have been accumulating with the 18 subjects using the GWP/MIT systems. Next quarter virtually all subjects will reach six months of use with that system. We will also begin longitudinal studies where parameters like stimulus rate, phase duration, and preemphasis will be altered in the eleven subjects not participating in the study of NLG mapping.

In the laboratory, we plan to revive our high-rate stimulation studies that were halted as we directed effort toward beginning the longitudinal testing program with the GWP/MIT system. Next quarter we will also test our first subject with binaural cochlear implants. Although this first test session of four days will mainly be devoted to characterizing the Ineraid and Clarion implants individually, we do expect to conduct some binaural stimulation.

Figure Captions

Figure 1. Single-syllable word and consonant identification scores for 18 subjects with Ineraid cochlear implants. These subjects used their commercial Ineraid sound processor at least four years before switching to the GWP/MIT sound processor. Each of these subjects was tested with the Auditec recording of the NU6, single-syllable word test (N=50) and the Iowa, 16-Consonant test (N=80) using their Ineraid and GWP/MIT sound processors. Black bars represent scores obtained with the subjects' Ineraid processor. In the case of Panel A, each black bar shows the average of the last three NU6 scores obtained before switching to the GWP/MIT system. In Panel B, black bars represent the Consonant test score obtained the day of switching. Moving right from the Ineraid bar, the open bars represent scores obtained with the GWP/MIT processor at 0, 1, 3 and 6 months postfitting where available. Subjects assigned numbers beginning with "S" were implanted at the Massachusetts Eye and Ear Infirmary by J.B. Nadol, Jr., M.D. Subject numbers beginning with "A" identify individuals implanted at other centers who joined our research program after implantation. The sound processors (both Ineraid and GWP/MIT) for all subjects were fitted by our research group.

Figure 2. Single-syllable word (panel A) and consonant identification (panel B) scores obtained with the GWP/MIT processor plotted as a function of the Ineraid score for each subject. Each symbol represents an average GWP/MIT score and an average Ineraid score for a single subject. The GWP/MIT averages include scores beginning at one month post fitting and beyond (see Figure 1). The Ineraid averages include the last three scores obtained with that processor before switching to the GWP/MIT system. Because all subjects used the Ineraid for at least four years before switching, the Ineraid average represents asymptotic performance.

Figure 3. Block diagram of the level mapping stage used in our laboratory and GWP/MIT sound processors. For each channel's center frequency, the total gain before the mapper is characterized by a single input gain (Gin). DR represents the output dynamic range. The mapper output amplitude modulates a pulse train that is converted to a current stimulus by the voltage-to-current source converter (V/I). Thus, the current pulse amplitude generated by a specific mapper input level is determined by the mapping function and Gout.

Figure 4. Comparison of loudness growth (magnitude estimates) for implant subject S01 using our standard laboratory sound processor with loudness growth for normal hearing subjects. Error bars correspond to plus and minus one standard deviation. In the case of the implantee, magnitude estimates were obtained for twenty levels of a 1300 Hz tone (center frequency for channel 3) presented to the external input of the processor. The twenty levels were logarithmically distributed over the 60 dB range shown in the figure. Each level was presented ten times in pseudo random order. The "randomization" of levels was accomplished by randomly drawing (without replacement) ten lists from a pool of twenty lists. The twenty lists (each of twenty levels) were selected from a larger

number of randomized lists to comply with the following constraints: each amplitude does not immediately precede any other amplitude more than twice and any two successive amplitudes do not differ by more than 60% of the linear dynamic range. A presentation consists of two 300 ms tones separated by approximately 300 ms. The subject was instructed as follows: "We will randomly present 20 sounds of different amplitudes. When the sounds are presented, a box drawn on the terminal 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). Answer "0" if you do not hear a sound. Do not worry about consistency. Simply try to match an appropriate number to each sound regardless of what you may have called the previous stimulus."

The dashed line depicting normal loudness growth was plotted using the relationship $L=kP^{0.6}$, where L represents loudness and P acoustic pressure. The acoustic level used as the reference (0 dB) produced loudness magnitude estimates in normal hearing subjects that correspond to those of the implant subject for electric stimulation levels at MCL.

Figure 5. Magnitude estimates obtained from subject S01 stimulating electrode 3 using twenty levels ranging from THR to MCL. The error bars represent plus and minus one standard deviation. The methods for magnitude estimation are the same as described in the caption for Figure 4. The biphasic, stimulus waveform (cathodic-first, 31.25 μs/phase, 2000 Hz, 5~~0~~ ms) was produced directly by controlling programmable current sources.
300

Figure 6. The top panel shows two mapping functions for S01's channel 3. The solid line represents the function designed to restore normal loudness growth and the broken line represents the standard logarithmic mapping function. The bottom panel plots magnitude estimates obtained from subject S01 using a 1300 Hz tone burst (center frequency for channel 3) for both the standard mapping function (open symbols) and the mapping function designed to restore normal loudness growth (closed symbols). The dashed line represents loudness growth for normal hearing listeners for 1 kHz acoustic tones.

References

Eddington DK, Delhorne L, Garcia N, Noel VA, Rabinowitz WM, Tierney J, Whearty ME (1996a): Speech Processors for Auditory Prostheses, Second Quarterly Progress Report. NIH Contract N01-DC-6-2100: #2.

Eddington DK, Tierney J, Rabinowitz WM, Noel VA, Whearty ME, Delhorne L, Garcia N (1996b): Speech Processors for Auditory Prostheses, First Quarterly Progress Report. NIH Contract N01-DC-6-2100: #1.

Fisher WM, Doddington GR, Goodie-Marshall KM (1986): The DARPA speech recognition research database: specifications and status: "DARPA Workshop on Speech Recognition." Palo Alto, California:, pp 93-99.

Tyler RS, Preece JP, Tye-Murray N (1987): Iowa Audiovisual Speech Perception Laser Videodisc Iowa City: The University of Iowa.

Wilson BS, Finley CC, Lawson DT, Wolford RD, Eddington DK, Rabinowitz WM (1991): Better speech recongition with cochlear implants. Nature 352:236-238.

Figure 1

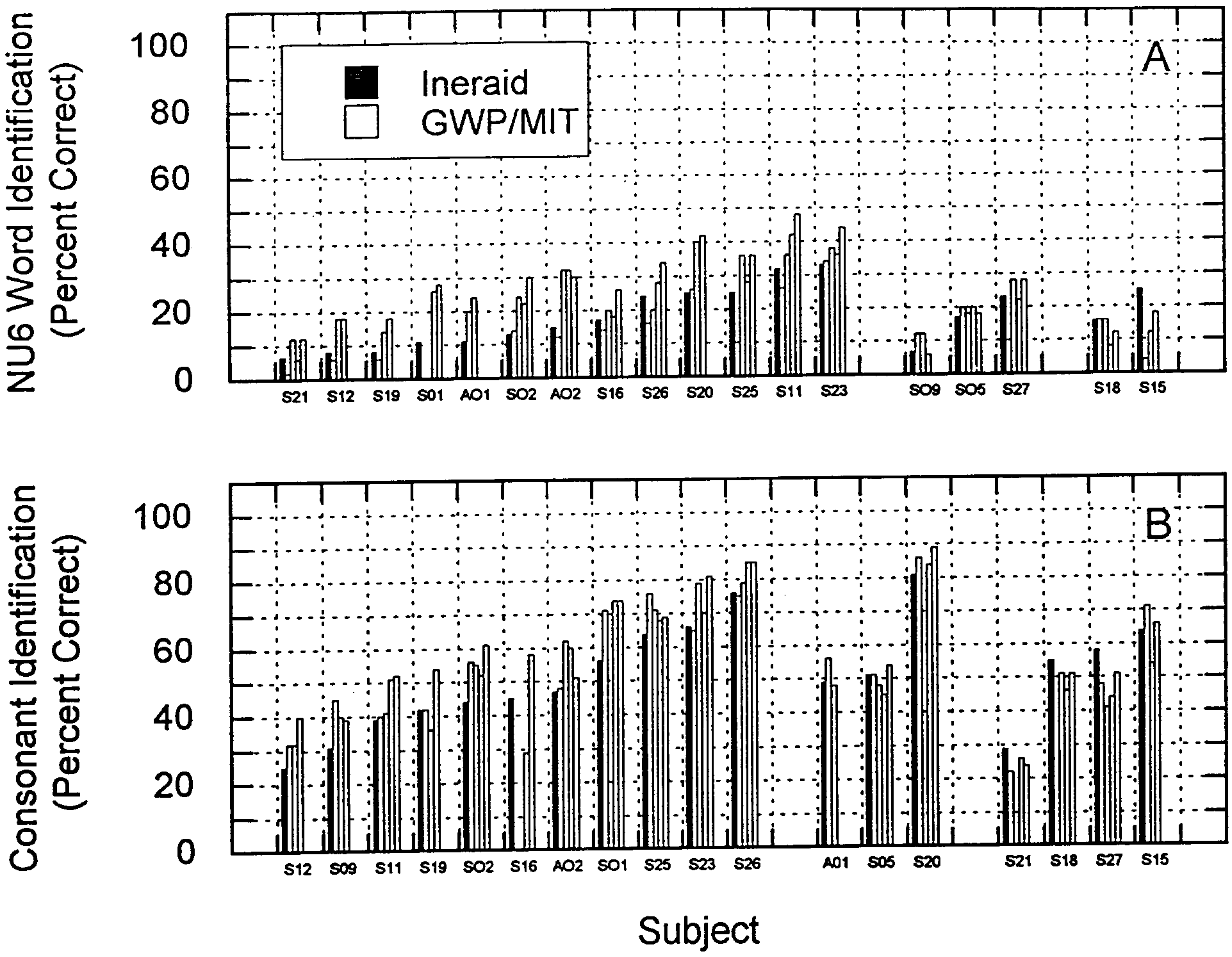
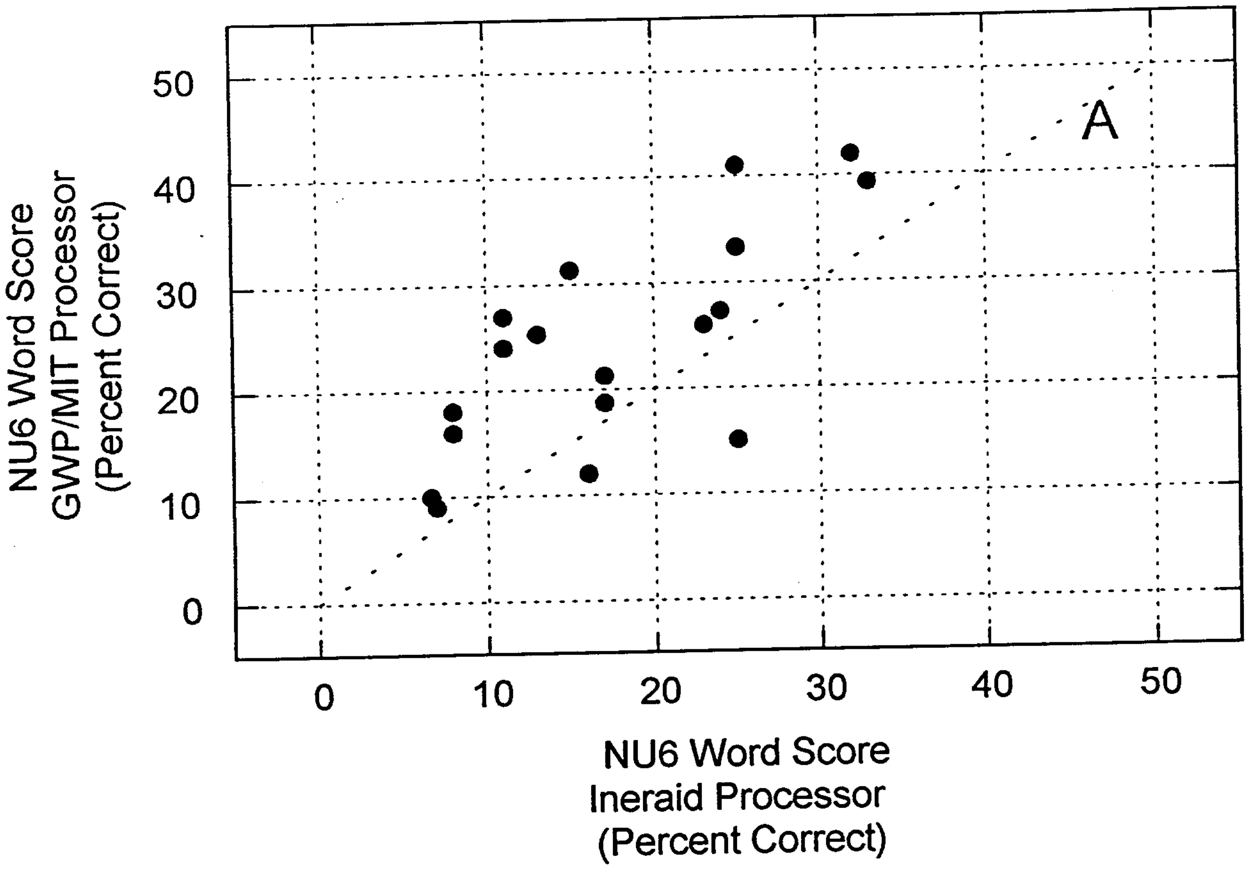
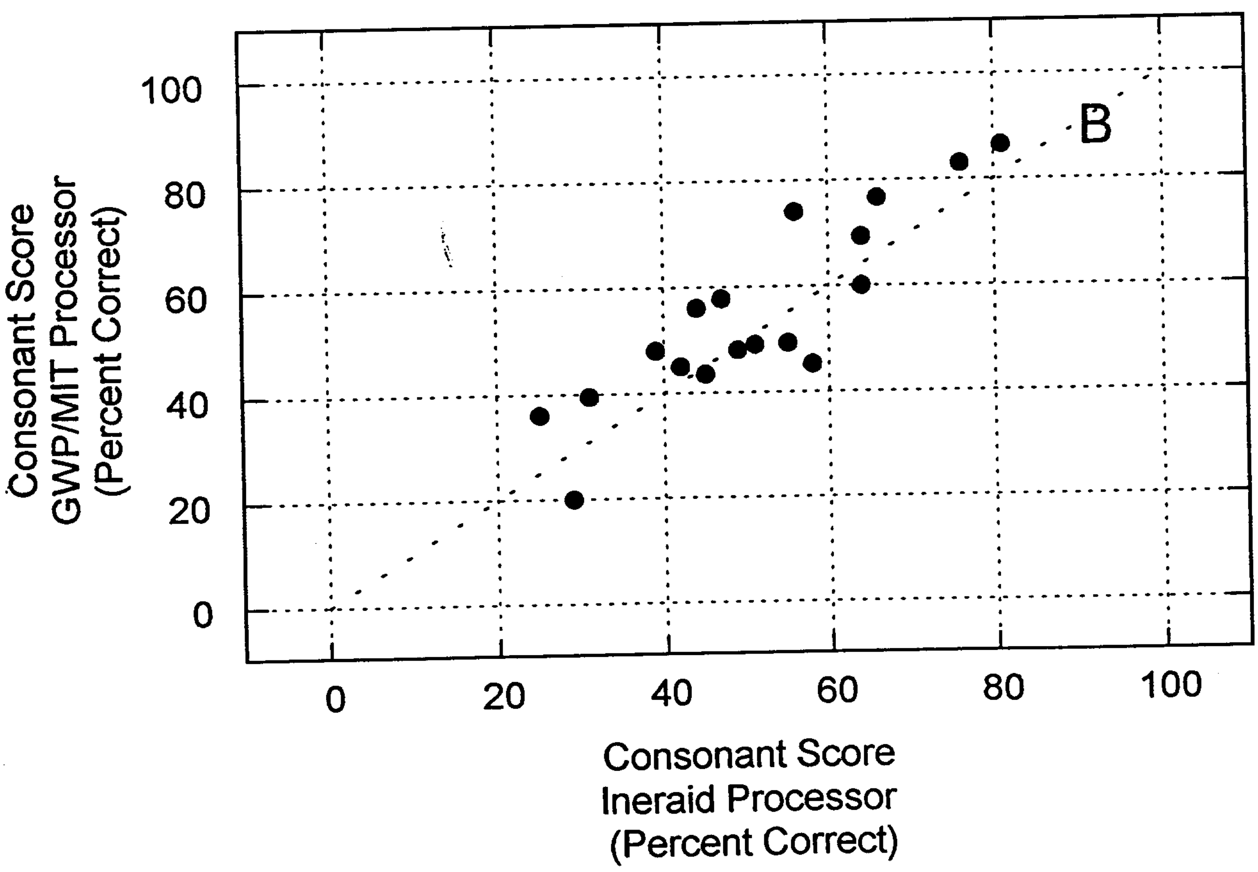

NU6 Word Identification (Percent Correct)
A
Ineraid
GWP/MIT
S21 S12 S19 S01 AO1 SO2 AO2 S16 S26 S20 S25 S11 S23 SO9 SO5 S27 S18 S15
Consonant Identification (Percent Correct)
B
S12 S09 S11 S19 SO2 S16 AO2 SO1 S25 S23 S26 A01 S05 S20 S21 S18 S27 S15
Subject

Figure 2


A
NU6 Word Score
GWP/MIT Processor
(Percent Correct)
0
10
20
30
40
50
NU6 Word Score
Ineraid Processor
(Percent Correct)


B
Consonant Score
GWP/MIT Processor
(Percent Correct)
0
20
40
60
80
100
Consonant Score
Ineraid Processor
(Percent Correct)

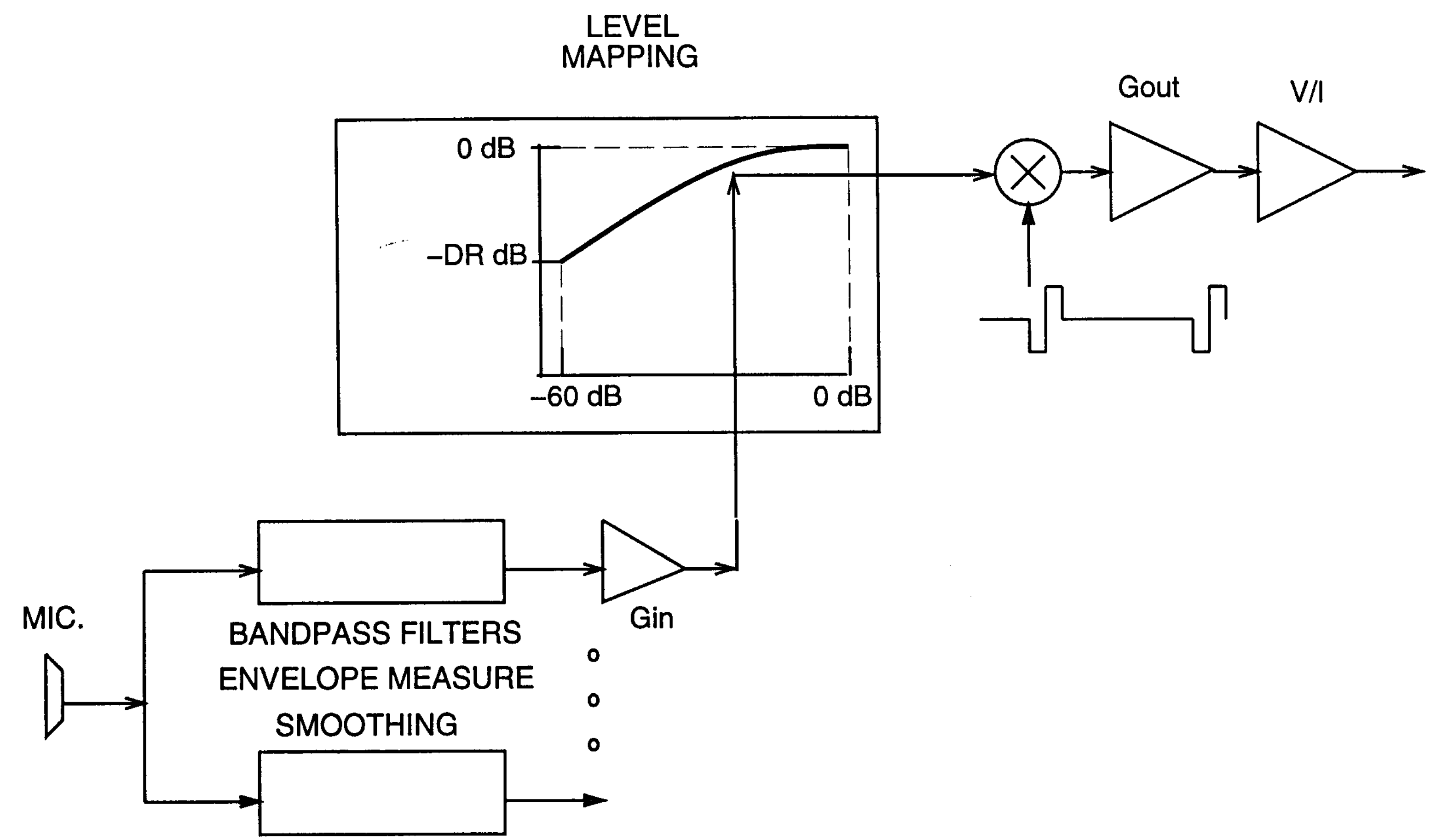
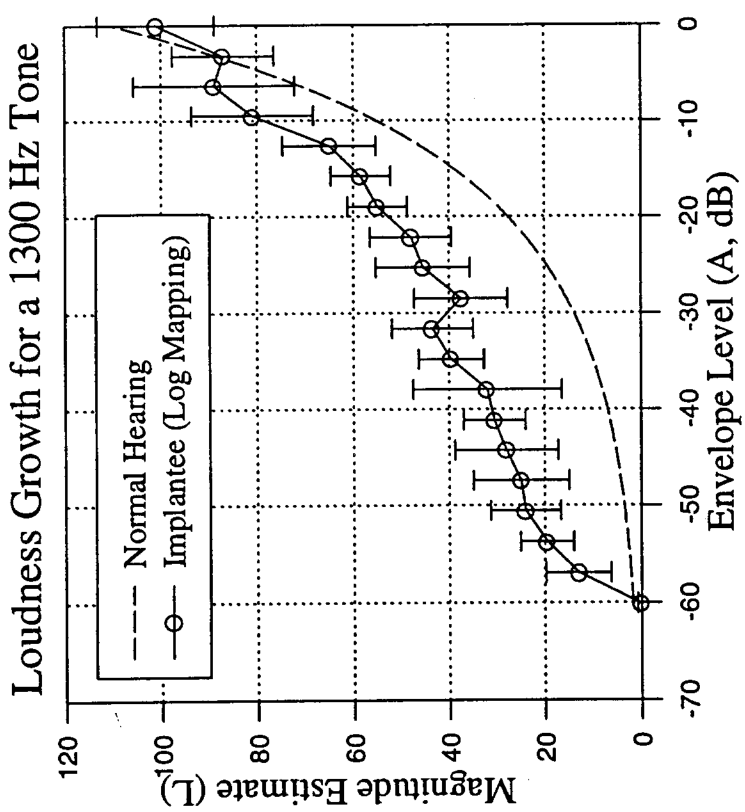

LEVEL
MAPPING
0 dB
-DR dB
-60 dB
0 dB
Gout
V/I
MIC.
BANDPASS FILTERS
ENVELOPE MEASURE
SMOOTHING
Gin

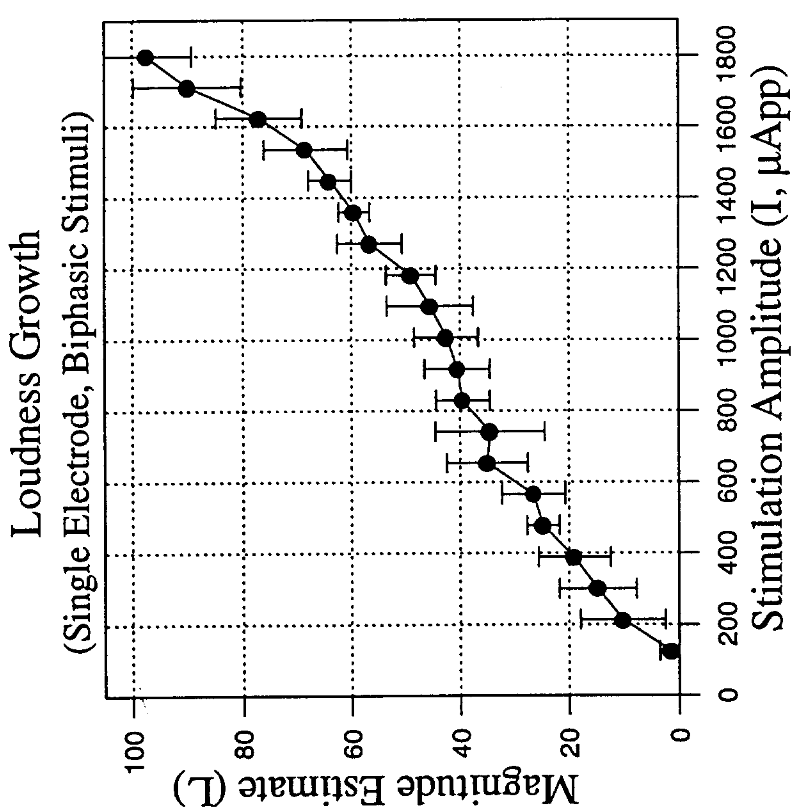
Figure 3

Figure 4



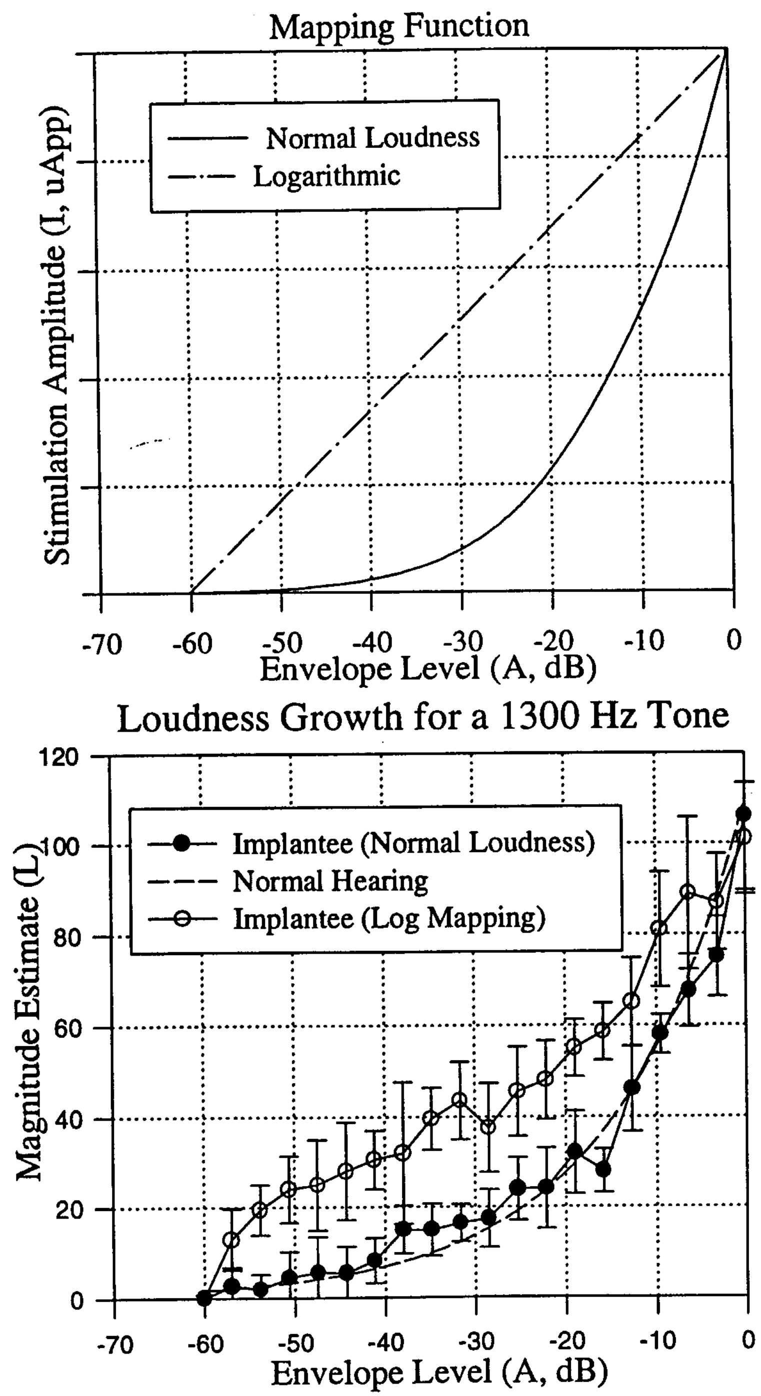
Loudness Growth for a 1300 Hz Tone
Normal Hearing
Implantee (Log Mapping)
Magnitude Estimate (L)
Envelope Level (A, dB)
120
100
80
60
40
20
0
-70
-60
-50
-40
-30
-20
-10
0

Figure 5



Loudness Growth
(Single Electrode, Biphasic Stimuli)
Magnitude Estimate (L)
0
20
40
60
80
100
Stimulation Amplitude (I, μApp)
0
200
400
600
800
1000
1200
1400
1600
1800

Figure 6


Mapping Function
Normal Loudness
Logarithmic
Stimulation Amplitude (I, uApp)
-70
-60
-50
-40
-30
-20
-10
0
Envelope Level (A, dB)
Loudness Growth for a 1300 Hz Tone
Implantee (Normal Loudness)
Normal Hearing
Implantee (Log Mapping)
Magnitude Estimate (L)
120
100
80
60
40
20
0
-70
-60
-50
-40
-30
-20
-10
0
Envelope Level (A, dB)