dup.

Fifth Quarterly Progress Report

January 1, 1997, through March 31, 1997

Speech Processors for Auditory Prostheses

NIH Contract N01-DC-6-2100

submitted by

Donald K. Eddington, PI
William M. Rabinowitz
Colette Boex-Spano
Joseph Tierney
Lorraine Delhorne
Nico Garcia
Victor A. Noel
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[1], 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

We are continuing to evaluate the speech-reception performance among subjects who have switched from their commercial Ineraid, compressed-analog sound processor to an alternative wearable unit that implements a version of the Continuous Interleaved Sampling (CIS) algorithm [1]. The CIS algorithm is implemented within the Geneva Wearable Processor (GWP) running a program developed by us (GWP/MIT) [2].

Seventeen subjects are now using the GWP/MIT systems on a regular basis. Fourteen subjects are being tested periodically as part of our longitudinal studies. Three others are tested as they are available but they are unable to comply with our formal testing schedule because of poor health and/or distance from Boston. We presented results from two speech reception tests (NU-6 and medial consonant) for these subjects in a previous QPR [3]. In this report, we present the results of four additional tests that have been conducted with these subjects.

A summary of results of five speech reception tests for the fourteen longitudinal subjects are given in Figure 1. Open-set tests include measures of isolated, monosyllabic word recognition (NU-6), recognition of words in everyday-like sentences (CUNY corpus), and word recognition in sentences against a background of speech-shaped noise (HINT

[1] Support for this work contributed by the Wilsdorf Foundation of Geneva.

corpus) presented at a speech-to-noise ratio of 10 dB. Closed-set identification of speech segments is assessed using sets of twelve initial consonants (12-CONS) and 10 medial vowels (10-VOWEL) from our own recordings. All subjects have now completed 6 months use of the GWP/MIT system. The results are given in terms of the mean and standard error obtained with the Ineraid processor and after 0 (day of fit), 1, 3, and 6 months use of the GWP/MIT system.

The different speech tests indicate similar patterns of change over time but their final outcomes are discrepant. All tests show an initial performance decrement (at the day of fit), a large growth in performance at 1 month, and smaller gains out to 3 and 6 months. However, while some tests show GWP/MIT performance at 6 months that substantially exceeds that with the Ineraid system, other tests indicate no such gains. The tests indicating gains are the open-set measures; particularly the CUNY sentences with means of Ineraid = 42% and GWP/MIT = 68%. Both closed-set tests indicate no gain.

Individual results on the CUNY sentences and our 12-consonant test further reinforce this apparent discrepancy. These results for the Ineraid system and with 6 months use of the GWP/MIT system are shown in Figure 2. Again, the general picture is that much larger gains are evident with the sentence test than with the consonant test. A few subjects show gains on both tests (S01 and S12) and two show decrements on both tests (S09 and S15). The other subjects, however, show essentially no change, or a decrement, on the consonant test but gains (of varying sizes) on the sentence tests. In some cases, the sentence gains are quite large (40% for S05, 54% for S26) while the corresponding consonant scores are unchanged.

A similar discrepancy was suggested from results in a previous report[4] where we compared results from a different consonant test (Iowa, medial-consonant) to those obtained using the NU-6, single-syllable word test. The general finding was that the closed-set, consonant test was less sensitive than the NU-6 test in documenting changes in performance.

These discrepancies between the results from different measures of performance are generally puzzling to us and merit further study. It is, however, comforting that (1) the strong subjective preference for the GWP/MIT system expressed by most subjects is, in general, consistent with the large gains in performance that are seen with some open-set tests and (2) it is these tests that have the strongest face validity in terms of relating to everyday communication.

3.0 Level Mapping: Magnitude Estimation of Loudness by Normal Hearing Subjects

In a previous QPR [3], we presented measures of the loudness growth function (LGF) for electric stimulation in Ineraid implantees. Using these measurements, we computed mapping functions for our standard interleaved processor that were designed to restore an approximation of normal loudness growth (NLG) for pure tones in these subjects. Because the process used to compute these mapping functions makes use of the equation $L = kP^{\alpha}$ (P: acoustic pressure of a tone) to describe the LGF for normal subjects, the value used for α will influence the shape of the mapping functions computed for the deaf, Ineraid subjects. In the previous report [3], we selected $\alpha = 0.6$ in order to make a few laboratory tests of NLG mapping functions in Ineraid subjects using the GWP/MIT implementation of interleaved processing.

Selecting a value for α from the literature is not straightforward. For instance, Stevens [5, 6] estimated α to be 0.6 based on the magnitude estimation data he obtained with normal hearing subjects that were instructed to make the ratios between the numbers they assigned to the various stimuli correspond to the ratios between the loudness of the sensations elicited by the stimuli. Hellman and Meiselman [7] also conducted magnitude estimation for loudness by instructing normal hearing subjects to assign numbers based on loudness but did not introduce concept of ratios. Their data were best fit by $\alpha = 0.46$.

Rather than selecting α from the literature, we decided to determine α by conducting magnitude estimation tasks in five normal-hearing subjects using the same procedure we used with the implantees to measure the LGF for single-electrode, electric stimulation. These measurements, especially as the number of normal-hearing subjects increases, should help us define the exponent to be used for characterizing normal-hearing LGFs when computing mapping functions designed to restore normal loudness growth in implantees.

3.1 Methods

Magnitude estimation was used to estimate the LGFs for five, normal-hearing subjects (three female, two male, ranging in age from 26 to 35 years). Measurements were made using two acoustic stimuli: (1) 1300 Hz tone and (2) speech-shaped noise (SSN) where the energy distribution across the noise bandwidth (150-7000 Hz) corresponded to the mean long-term averaged speech spectrum [8]. All measurements were conducted monaurally, using TDH 39 headphones.

Magnitude estimates were obtained for twenty levels of the tone and twenty levels of the noise. These levels ranged between the threshold (THR) and most comfortable loudness (MCL) levels measured for both tone and SSN in each subject in the same way THR and MCL are measured in our implant subjects[2, 3]. Based on preliminary experiments and because our subject instructions are similar to those of Hellman and Meiselman [7], we distributed the twenty levels to produce a uniform distribution in loudness based on the relationship $L=cP^{0.46}$. 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 approximately 80% of the difference between THR and MCL.

A presentation consisted of two, 300 ms signals of the same amplitude (rise-fall time: 10 ms) separated by approximately 400 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."

3.2 Results

Tables 1 and 2 show the THR and MCL levels measured for the five normal-hearing subjects for the 1300 Hz tone and the SSN, respectively. Thresholds are significantly higher for the SSN than for the 1300 Hz tone (11.8 dB SPL (SD: 0.84) vs. 4.6 dB SPL (SD: 1.52), $P < 1.10^{-4}$), MCLs are not statistically different (80.2 dB SPL (SD: 10.03) vs. 82.4 dB SPL (SD: 2.7), $P > 0.25$). Note that all sound level measurements were C-weighted and made using a 6cc headphone coupler.

Table 1
Psychophysics for a 1300 Hz Tone Burst

Subject	THR (dB SPL)	MCL (dB SPL)
NH1	6	82
NH2	6	79
NH3	3	81
NH4	5	86
NH5	3	84

Table 2
Psychophysics for Speech-Shaped Noise

Subject	THR (dB SPL)	MCL (dB SPL)
NH1	13	73
NH2	11	85
NH3	12	84
NH4	12	92
NH5	11	67

Figure 3 shows loudness growth measurements for subject NH3 made with the 1300 Hz tone and the SSN. Notice the least-squares fit to the tone and noise data shown in the right panels. The slopes of these two lines were used to compute the exponent, α, ($\alpha = 20 \cdot$ slope) for the power functions associated with these data. As shown in the bottom-right panel, the straight-line fit of the noise data was applied to signal levels greater than 40 dB SPL (well above threshold).

Table 3 lists the exponents computed for the five normal-hearing subjects for both the tone and SSN stimuli. The difference between the α_{mean} computed from our 1300 Hz tone data and the α_{mean} for the data of Hellman and Meiselman is small (0.04), but statistically significant (0.46 (SD: 0.02), N=51; 0.42 (SD: 0.09), N=5; $P<0.005$). Given the small number of subjects and the large range of the individual αs, it is not surprising that the

Table 3 α for $L=kP^{\alpha}$		
Subject	1300 Hz	SSN
NH1	0.38	0.27
NH2	0.32	0.28
NH3	0.44	0.33
NH4	0.55	0.54
NH5	0.39	0.40

difference between the α_{mean} for the LGFs obtained using a 1300 Hz tone and the α_{mean} for the SSN are not statistically significant (α_{mean} =0.42 (SD: 0.09), N=5; α_{mean} =0.36 (SD: 0.11), N=5; $P > 0.4$). The range of αs measured using 1300 Hz tone bursts is consistent with the range described by Hellman and Meiselman for their 51 subjects measured using 1000 Hz tone bursts.

The relatively large range of individual αs shown in Table 3 raises several questions regarding the use of a single α in computing mapping functions designed to produce normal loudness growth across a population of implantees. One might wonder whether the range of αs across subjects is due to the measurement process or whether they represent real, individual differences. While the preliminary nature of our data (small N; few within-subject, repeated measures) limit its power in addressing this question, the correlation between αs for 1300 Hz tone bursts and SSN (r=0.87) is consistent with evidence from a number of studies showing that differences between individual αs are repeatable and stable with time (e.g., [9]). This means that if subject performance benefits from mapping functions that restore the normal growth of loudness, α may require adjustment for each subject in order to compute NLG maps that provide optimum performance.

4.0 Future Work

Next Quarter we will continue work directed at normal loudness growth (NLG) mapping functions. Using the GWP/MIT device that allows us to easily switch between mapping functions, we will investigate whether subjects are able to discriminate three NLG mapping functions based on three αs that represent the range shown in Table 3 (0.27, 0.42, 0.55). If subjects are able to discriminate between the three mapping functions, we will also determine whether they are able to consistently rank order them by preference. We also plan to conduct three to four-week trials to evaluate speech-reception performance for these NLG mapping functions.

We are continuing the analysis of the longitudinal speech-reception results accumulated with the 18 subjects using the GWP/MIT systems. These analyses will focus on relationships between the different speech-reception tests. We will also continue longitudinal studies where parameters like electrode configuration, stimulus rate, phase duration, and preemphasis will be altered in subjects not participating in the study of NLG mapping.

In the laboratory, we plan to conduct a second battery of tests with our first, bilaterally implanted subject.

Figure Captions

Figure 1. Results of five speech reception tests for the fourteen Ineraid subjects as they switched from their compressed-analog Ineraid processor to a wearable system that implements a form of continuous interleaved sampling (CIS). The results are grouped by test (see text for test descriptions) and the height of each bar represents the mean score obtained by the fourteen subjects for a particular speech-reception test and time. The first bar of each group shows mean performance of the subjects using the Ineraid processor (each subject had at least two years of experience with this processor). The other bars of each group are the mean scores obtained at different times relative to the day of fitting and show performance as a function of experience with the new processing system. Error bars mark one standard error.

Figure 2. The top panel presents CUNY sentence scores for fourteen subjects. The bottom panel shows scores for the 12-(initial)consonant test. The shaded bars represent Ineraid scores and the dark bars show the CIS scores at 6 months.

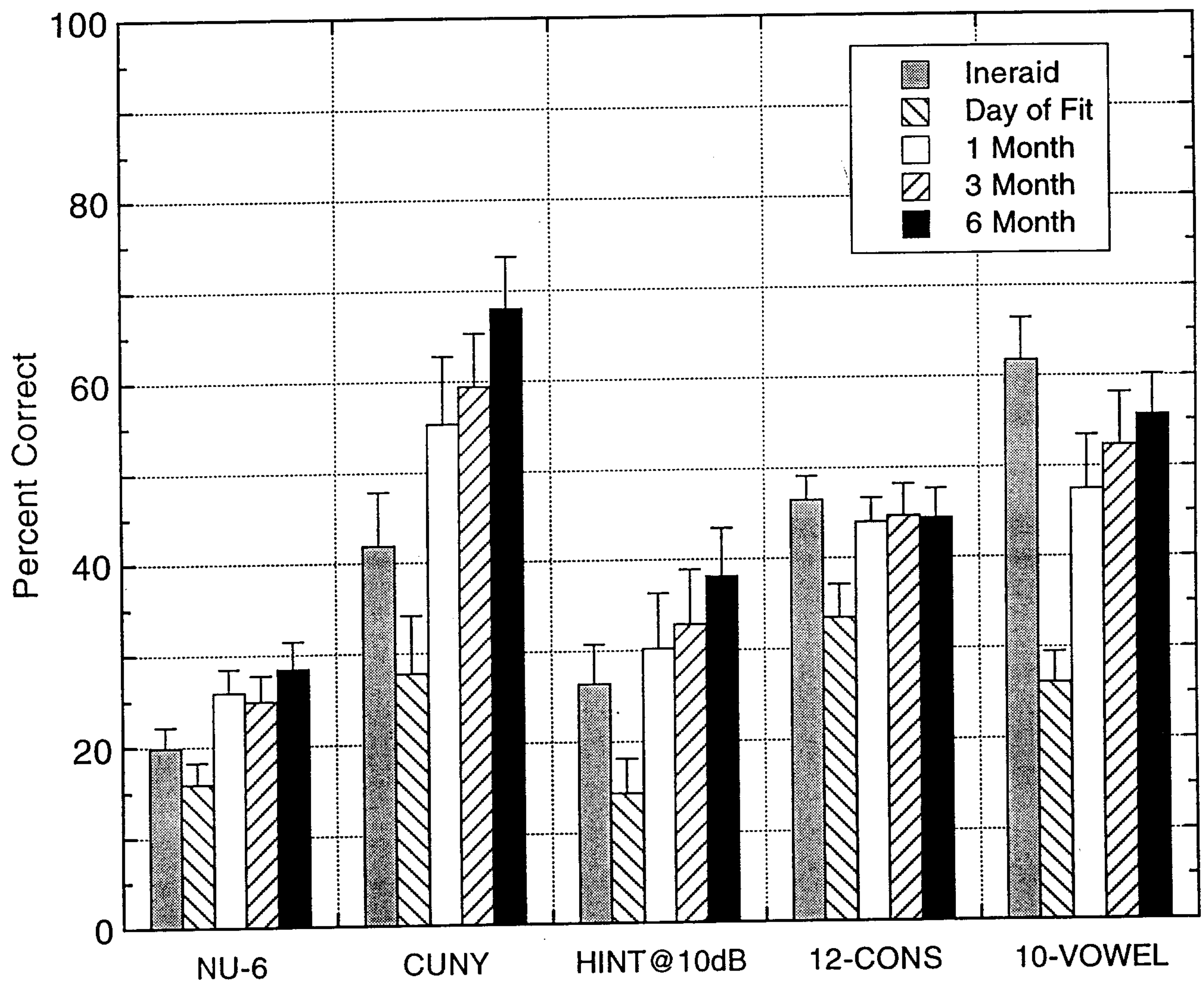
Figure 3. Magnitude estimates made by the normal-hearing subject NH3 plotted as a function of the acoustic stimulus level. Results for the 1300 Hz tone are shown in the top panels and for the SSN in the bottom panels. The magnitude estimates are represented linearly in the left panels and logarithmically in the right panels. Error bars represent interquartiles. Solid lines represent growth functions computed from least-squares, straight-line fits to the data represented by the left panels.

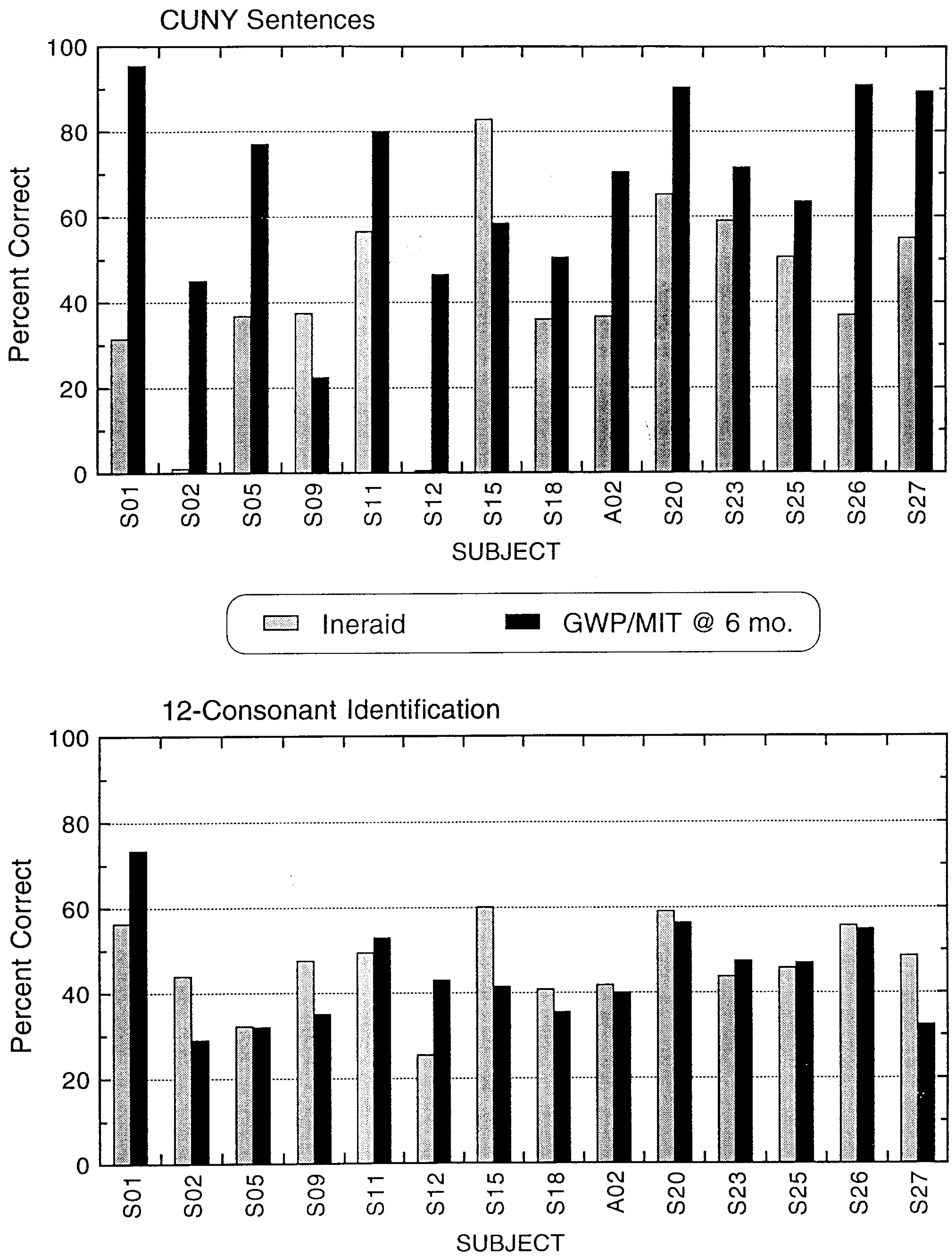
Bibliography

1. B.S. Wilson, C.C. Finley, D.T. Lawson, R.D. Wolford, D.K. Eddington, and W.M. Rabinowitz, "Better speech recongition with cochlear implants," Nature . 352: p. 236-238 (1991).
2. D.K. Eddington, J. Tierney, W.M. Rabinowitz, V.A. Noel, M.E. Whearty, L. Delhorne, and N. Garcia, Speech Processors for Auditory Prostheses, First Quarterly Progress Report, (Cambridge: MIT Research Laboratory of Electronics, 1996).
3. D.K. Eddington, C. Boex-Spano, L. Delhorne, N. Garcia, V.A. Noel, W.M. Rabinowitz, J. Tierney, and M.E. Whearty, Speech Processors for Auditory Prostheses, Third Quarterly Progress Report, (Cambridge: MIT Research Laboratory of Electronics, 1996).
4. D.K. Eddington, L. Delhorne, N. Garcia, V.A. Noel, W.M. Rabinowitz, J. Tierney, and M.E. Whearty, Speech Processors for Auditory Prostheses, Second Quarterly Progress Report, (Cambridge: MIT Research Laboratory of Electronics, 1996).
5. S.S. Stevens, "The measurment of loudness," Journal of the Acoustical Society of America . 27(5): p. 815-829 (1955).
6. S.S. Stevens, "Concerning the form of the loudness function," Journal of the Acoustical Society of America . 29: p. 603-606 (1957).
7. R.P. Hellman and C.H. Meiselman, "Prediction of individual loudness exponents from cross-modality matching," Journal of Speech and Hearing Research . 31: p. 605-615 (1988).
8. D. Byrne, *et al.*, "An international comparison of long-term average speech spectra," Journal of the Acoustical Society of America . 96(4): p. 2108-2120 (1994).
9. R.P. Hellman, "Stability of individual loudness functions obtained by magnitude estimation and production," Perception and Psychophysics . 29: p. 63-70 (1981).

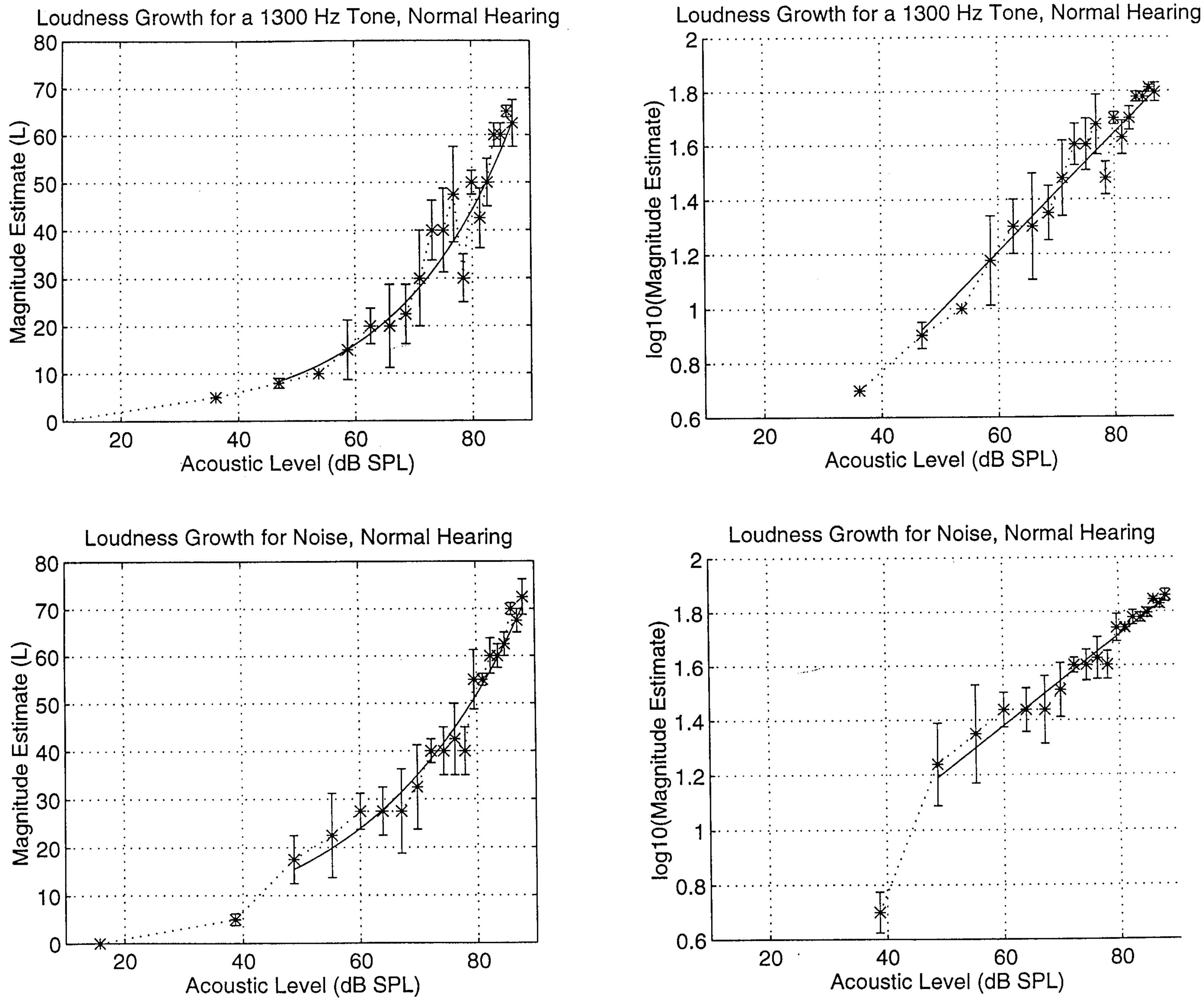
Ineraid → GWP/MIT

(mean and +1 s.e., n=14)





CUNY Sentences
Percent Correct
100
80
60
40
20
0
S01
S02
S05
S09
S11
S12
S15
S18
A02
S20
S23
S25
S26
S27
SUBJECT
Ineraid
GWP/MIT @ 6 mo.
12-Consonant Identification
Percent Correct
100
80
60
40
20
0
S01
S02
S05
S09
S11
S12
S15
S18
A02
S20
S23
S25
S26
S27
SUBJECT


Loudness Growth for a 1300 Hz Tone, Normal Hearing
Magnitude Estimate (L)
Acoustic Level (dB SPL)
Loudness Growth for a 1300 Hz Tone, Normal Hearing
log10(Magnitude Estimate)
Acoustic Level (dB SPL)
Loudness Growth for Noise, Normal Hearing
Magnitude Estimate (L)
Acoustic Level (dB SPL)
Loudness Growth for Noise, Normal Hearing
log10(Magnitude Estimate)
Acoustic Level (dB SPL)