Seventh Quarterly Progress Report

July 1, 1997, through September 30, 1997

Speech Processors for Auditory Prostheses

NIH Contract N01-DC-6-2100

submitted by

Donald K. Eddington
Colette Boex-Spano
Joseph Tierney
Victor Noel
Margaret Whearty

Massachusetts Institute of Technology
Research Laboratory of Electronics
Cambridge, MA

This QPR is being sent to
you before it has been
reviewed by the staff of the
Neural Prosthesis Program.

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.

The material of this report relates to 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.

The initial goal of the work we report this Quarter was to design and test CIS mapping functions that restore normal loudness growth (NLG) of tones for implantees. This work also led us to revisit methods used to define the maximum stimulation currents (I_{maxs}) 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) loudness growth (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".

---

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

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.

2.0 Methods

2.1 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 processor (Eddington et al., 1994), 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 MIT (GWP/MIT, Eddington et al., 1995).

2.2 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 μs/ph, except for subject S05, where 40 μs 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 detected (Levitt, 1970). 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):

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

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.

2.3 Sound Processors

We used the laboratory digital signal processing facility as described in a previous report (Eddington et al., 1995) 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 can also be found in a previous report (Eddington et al., 1995). The strategy is shown in Figure 1 as a block-diagram that emphasizes the mapping function. A range of acoustic frequencies from 250 Hz to 7000 Hz, are divided into 6 logarithmically segmented frequency bands. The output of each of these bands is passed to an envelope extractor (using a quadrature technique) and followed by a lowpass smoothing filter (400 Hz) The range of band envelope levels is then "compressed" by a function that maps the range of envelope levels to the electric dynamic range of the band's corresponding, intracochlear electrode. Finally, each intracochlear electrode is stimulated by a biphasic pulse train, that is amplitude modulated by the compressed envelope signal from its respective band. Biphasic pulses (cathodic first, 31.25 µs/phase, 2000 Hz) are used to stimulate electrodes and they are interleaved across electrodes to guarantee that only one electrode is stimulated at a time.

2.4 α and L Mapping functions

At the input of the mapping function, a gain (G_{in}, see Figure 1) 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 are 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 2 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_{maxs} obtained by this method are used to define the logarithmic mapping functions called "L functions".

2.5 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:

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

where P is the acoustic pressure of a tone. The exponent α has been defined by Stevens (1955, 1957) 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 conducted 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:

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

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:

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

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

2.6 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 (Tyler, Preece and Tye-Murray, 1987).

3.0 Results

3.1 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_{maxs} are based on an equal loudness criterion) for a few weeks. The left panel in Figure 3 shows scores of 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. The scores are expressed in percent correct. 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 3 describes 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 words scores are consistent with the clear preference articulated by all subjects for the L logarithmic mapping function processors.

3.2 Loudness growth

Figure 4 shows 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 ($L_e(I_e) = a_{0,e} + a_{1,e} \cdot I_e + a_{2,e} \cdot I_e^2 + a_{3,e} \cdot I_e^3 + a_{4,e} I_e^4$) 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 5 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 6 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 stimulation (Zeng and Shannon, 1992). Some of the growth functions of S15, however, are much more similar to the exponential.

3.3 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 7 plots 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 1 shows the exponent preferred by each subject (when more than one was tested) and, in some cases, comments relating to their preference.

TABLE 1

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

3.4 Speech reception with NLG mappings

While the new mapping functions are designed to restore the normal growth of loudness for single tones, we wanted to evaluate the effect of these mapping functions on speech reception. To date, five of the seven subjects for whom NLG mapping functions have been determined have used CIS systems based on these mapping functions for more than three weeks. Figure 8 shows the mean identification scores computed for at least 10 randomized presentations of the 24 (or 16 for subject S05) consonants for these five subjects. 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 8 shows scores for single-syllable word recognition. Notice that the two subjects (S01 and S02) who demonstrate 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).

4.0 FUTURE WORK

We expect that much of the next quarter will be directed at preparing a competitive renewal application for a continuation of this Contract.

In addition, we also expect to complete several experiments in normal listeners using acoustic simulations of CIS processing. These experiments are directed at estimating whether the current cochlear implant subjects using CIS processing are extracting the majority of the information presented.

We also expect to finish testing at least two Boston subjects with Geneva/GWP processors that were fit by investigators from Geneva. In the longer term, we plan to fit the same two Boston subjects with RTI/GWP processors and also plan to fit at least two Geneva subjects with MIT/GWP processors. The speech reception results from these experiments should help to determine whether any of the differences between these CIS processing schemes are functionally important.

References

Eddington DK, Rabinowitz WM, Tierney J, and Zissman MA (1994). Speech processors for auditory prostheses. Quarterly Progress Report (July 1, 1994 through September 30, 1994). NIH Contract N01-DC-2-2402.

Eddington DK, Noel VA, Rabinowitz WM, Svirsky MA, Tierney J, and Zissman MA (1995). Speech processors for auditory prostheses. Final report (September 30, 1992 through December 31, 1995). NIH Contract N01-DC-2-2402.

Eddington DK, Rabinowitz WM, Boëx-Spano C, Tierney J, Delhorne LA, Garcia N, Noel VA and Whearty ME (1997). Speech processors for auditory prostheses. Fifth Quarterly Progress Report (January 1, 1997, through March 31, 1997). NIH Contract N01-DC-6-2100.

Fisher WM, Doddington FG and Goodie-Marshall KM (1986). The DARPA speech recognition research database: specifications and status: "DARPA workshop on speech recognition" Palo Alto, California, p: 93-99.

Hellman RP and Meiselman CH (1988). Prediction of individual loudness exponents from cross-modality matching. J.Speech and Hearing Res. 31: 605-615.

Levitt H (1970). Transformed up-down methods in psychophysics. J. Acoust. Soc. Am. 49: 467-477.

Stevens SS (1955). The measurements of loudness. J. Acous. Soc. Am. 27(5): 815-29.

Stevens SS (1957). Concerning the form of the loudness function. J. Acoust. Soc. Am. 29: 603-6.

Tyler RS, Preece JP and Tye-Murray N (1987). Iowa audiovisual speech perception laser videodisc Iowa City: The University of Iowa.

Zeng FG and Shannon RR (1992). Loudness balance between electric and acoustic stimulation. Hear. Res. 60: 231-5.

Figure Captions

Figure 1. Block diagram of a CIS processor that emphasizes the stages associated with level mapping. For each channel, 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 2. 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 3. 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.

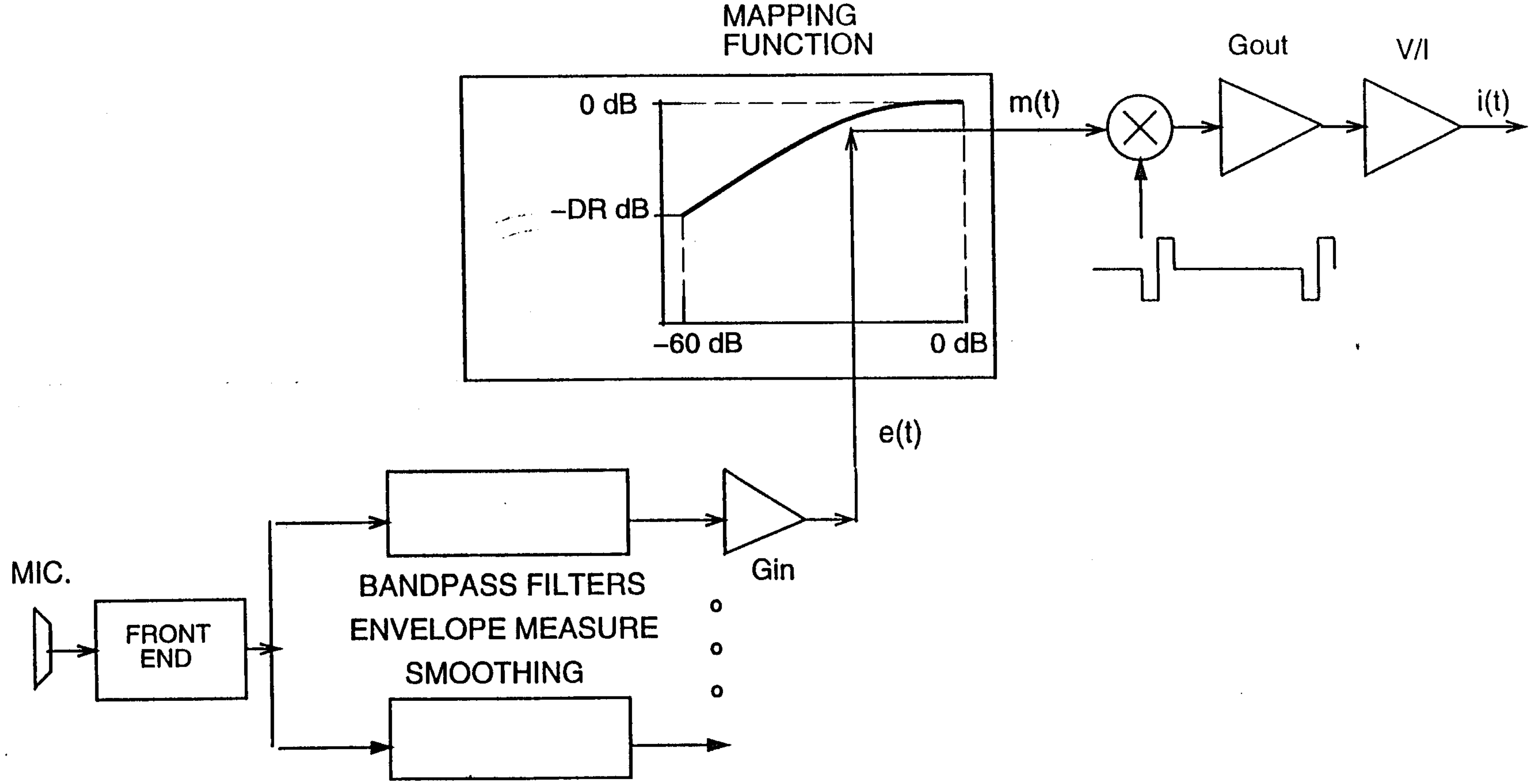
Figure 4. Subject S01's most apical electrode's loudness growth function. The solid line represents the 4th order polynomial fitted loudness growth function $L_e(I_e)$. The error bars represent the interquartiles of the 8 magnitude estimates obtained for each of the 20 different amplitudes evenly distributed 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.

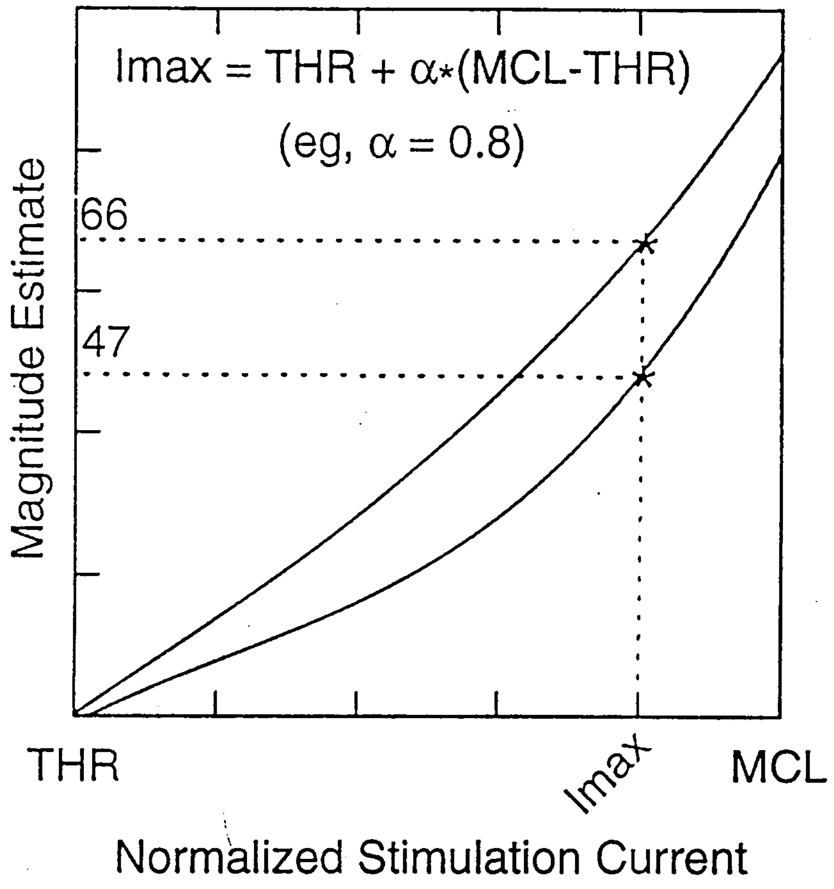
Figure 5. Loudness Growth functions obtained for the seven subjects for each of their electrodes. The solid lines represent the 4th order polynomial fitted loudness growths $L_e(I_e)$ defined from magnitude estimates realized between the THR and the MCL for each electrode.

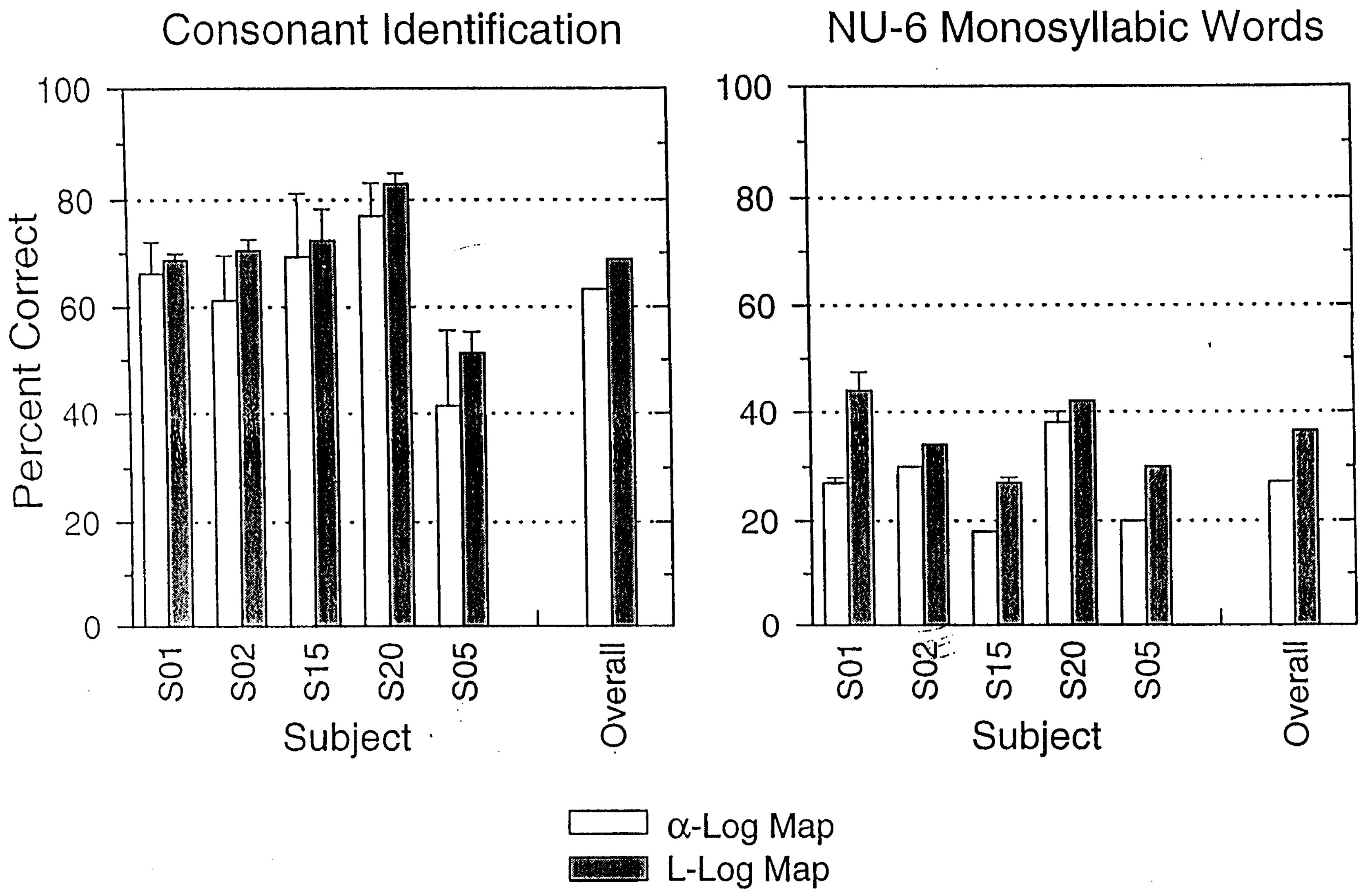
Figure 6. Loudness Growth functions obtained for each electrode of subject S01 (left panel) and of subject S15 (right panel). The solid lines are the fits to the AME results and the dashed lines are examples of exponential growth functions over the same range.

Figure 7. 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.

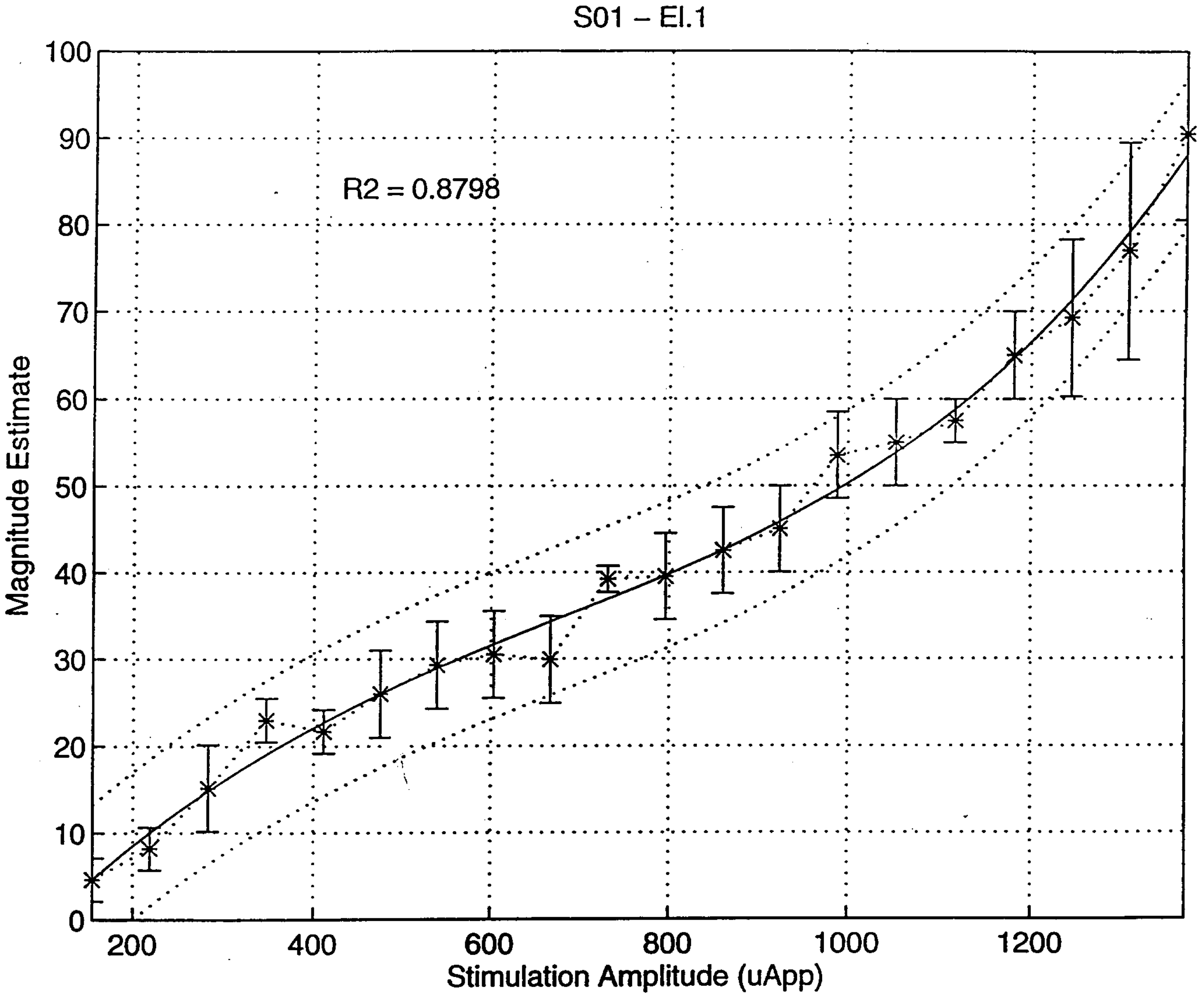
Figure 8. The left panel shows consonant identification performance for CIS processors using L logarithmic mapping functions (gray) and NLG mapping functions (white). Mean consonant identification scores are computed using at least 10 randomized presentations of the consonant lists for each subject. The right panel presents results for single-syllable word recognition.

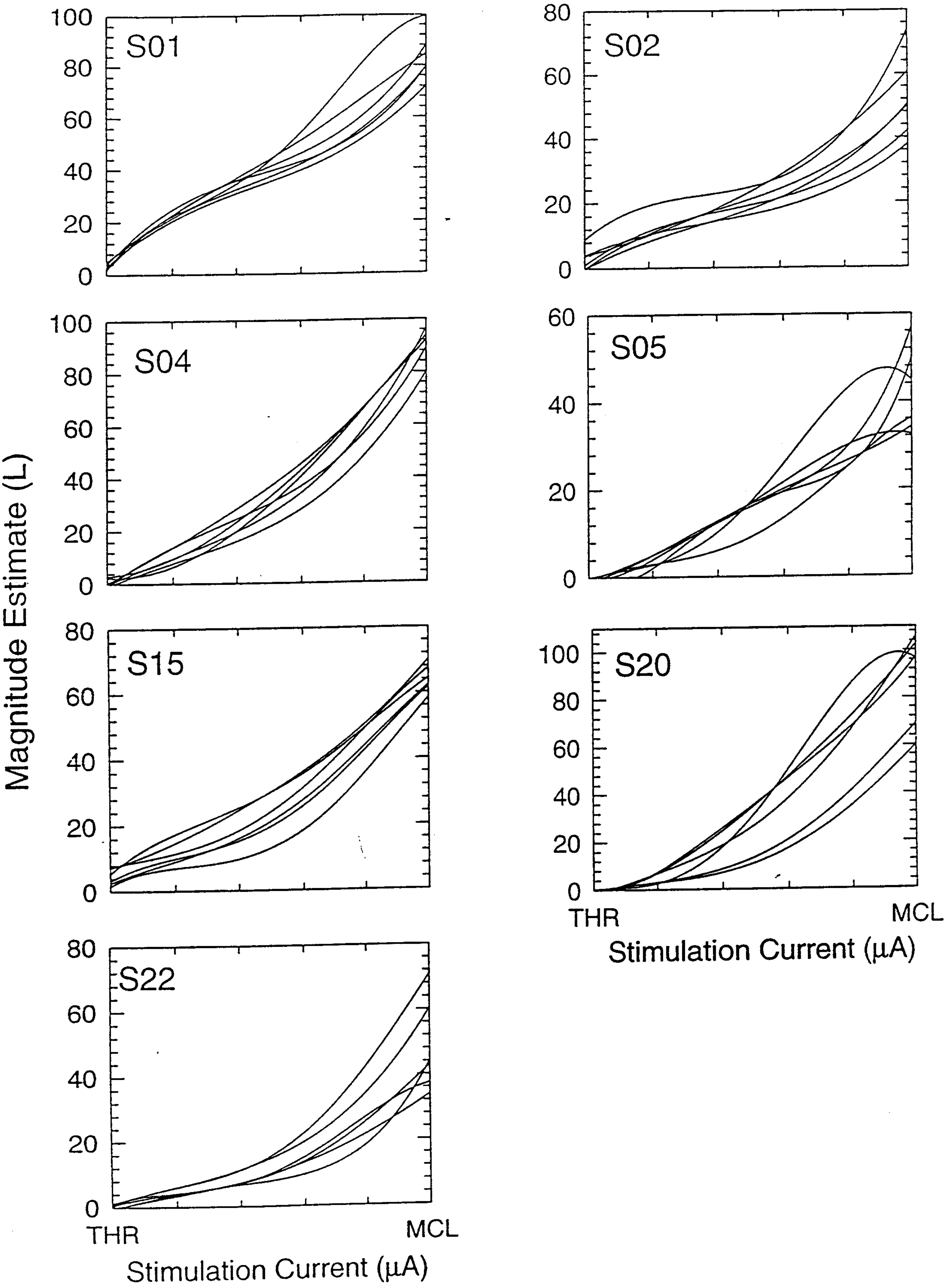

MAPPING
FUNCTION
0 dB
–DR dB
–60 dB
0 dB
m(t)
Gout
V/I
i(t)
e(t)
Gin
MIC.
FRONT END
BANDPASS FILTERS
ENVELOPE MEASURE
SMOOTHING

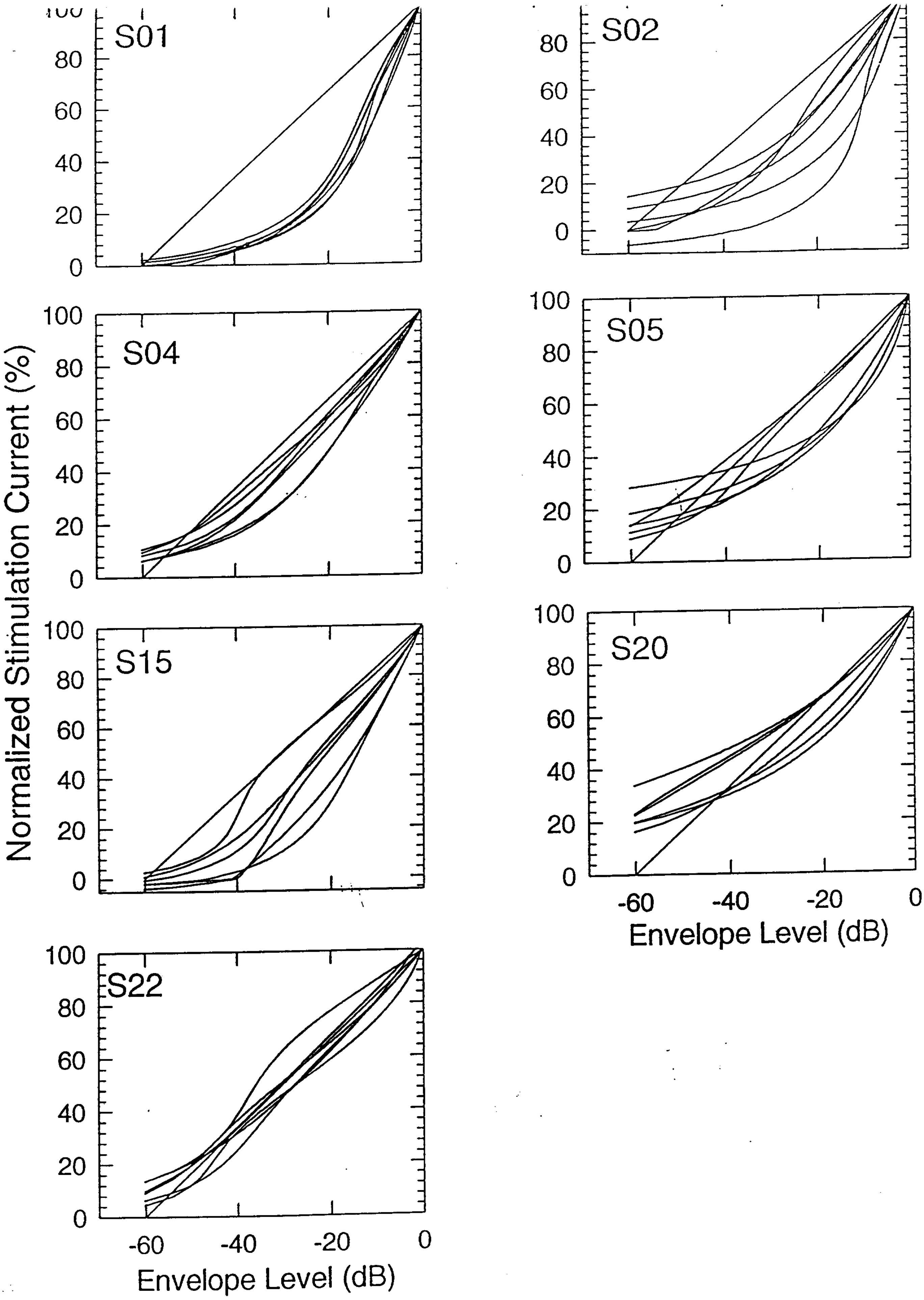

Imax = THR + α*(MCL-THR)
(eg, α = 0.8)
Magnitude Estimate
66
47
THR
Imax
MCL
Normalized Stimulation Current

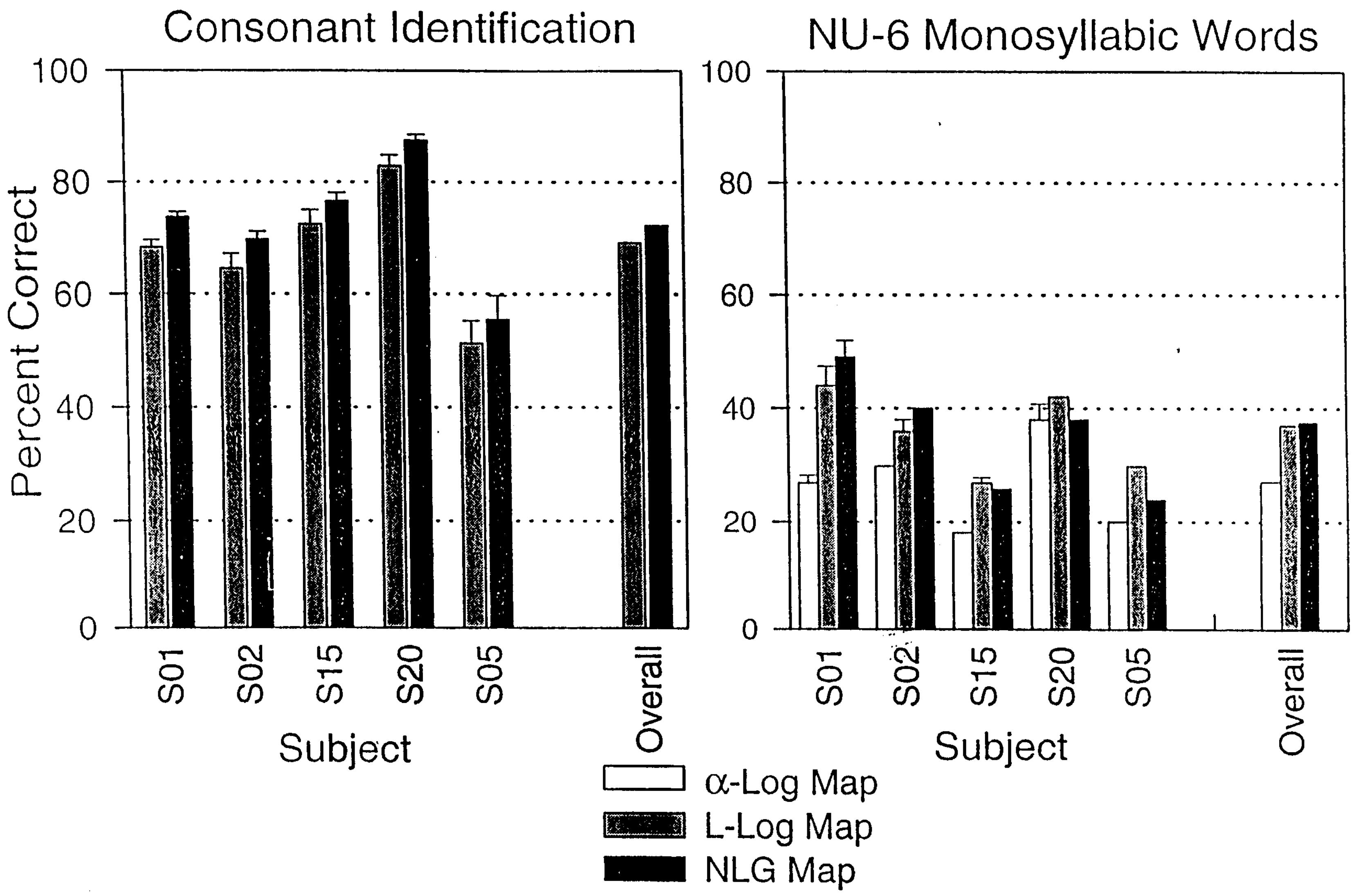

Consonant Identification

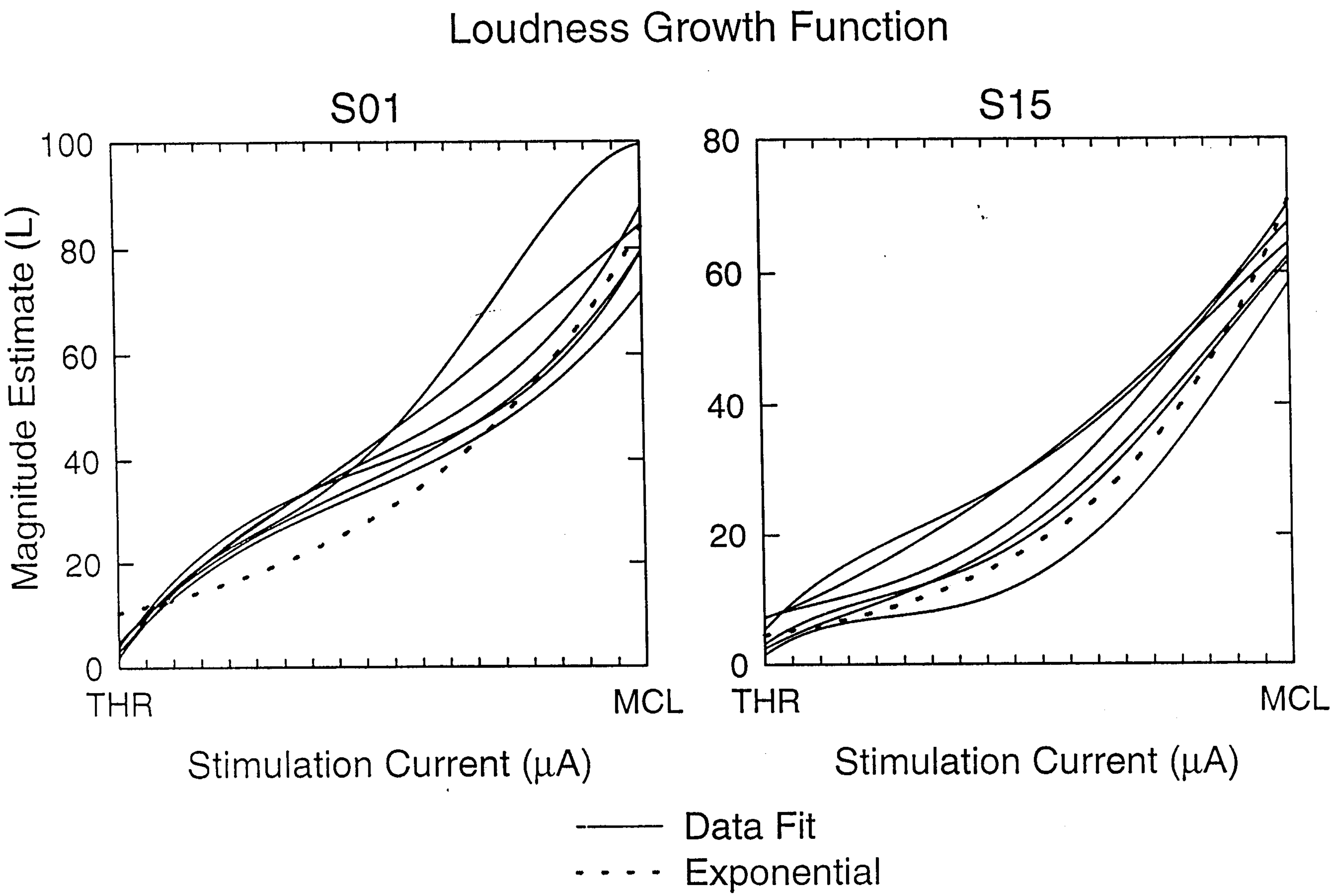
NU-6 Monosyllabic Words


S01 – El.1
R2 = 0.8798
Magnitude Estimate
Stimulation Amplitude (uApp)
0
10
20
30
40
50
60
70
80
90
100
200
400
600
800
1000
1200


S01
S02
S04
S05
S15
S20
S22
Magnitude Estimate (L)
0
20
40
60
80
100
THR
MCL
Stimulation Current (μA)


S01
S02
S04
S05
S15
S20
S22
Normalized Stimulation Current (%)
100
80
60
40
20
0
-60
-40
-20
0
Envelope Level (dB)


Consonant Identification
NU-6 Monosyllabic Words
Percent Correct
0
20
40
60
80
100
S01
S02
S15
S20
S05
Overall
Subject
α-Log Map
L-Log Map
NLG Map


Loudness Growth Function
S01
S15
Magnitude Estimate (L)
100
80
60
40
20
0
80
60
40
20
0
THR
MCL
THR
MCL
Stimulation Current (μA)
Stimulation Current (μA)
Data Fit
Exponential