REVISED

The effect of reduced dynamic range on speech understanding: Implications for patients with cochlear implants

Philipos C. Loizou
Department of Electrical Engineering
University of Texas at Dallas
Richardson, TX 75083-0688

Michael Dorman
Department of Speech and Hearing Science
Arizona State University, Tempe, AZ 85287-0102
and
University of Utah Health Sciences Center
Salt Lake City, Utah 84132

Jeanette Fitzke
Department of Speech and Hearing Science
Arizona State University
Tempe, AZ 85287-0102

Address correspondence to:
Philipos C. Loizou, Ph.D.
Department of Electrical Engineering
Univ. of Texas at Dallas
P.O.Box 830688, EC 33
Richardson, TX 75083-0688
E-mail: loizou@utdallas.edu

Phone: (972) 883-4617
Fax: (972) 883-2710

Objective: To determine the effect of reduced dynamic range on speech understanding when the speech signals are processed in a manner similar to a 6-channel cochlear-implant speech processor.

Design: Signals were processed in a manner similar to a 6-channel cochlear implant processor and output as a sum of sine waves with frequencies equal to the center frequencies of the analysis filters. The amplitudes of the sine waves were compressed in a systematic fashion to simulate the effect of reduced dynamic range. The compressed signals were presented to 10 normal-hearing listeners for identification.

Results: There was a significant effect of compression for all test materials. The effect of the compression on speech understanding was different for the three test materials (vowels, consonants and sentences). Vowel recognition was affected the most by the compression, while consonant recognition was affected the least by the compression. Feature analysis indicated that the reception of place information was affected the most. Sentence recognition was moderately affected by the compression.

Conclusions: Dynamic range should affect the speech perception abilities of cochlear implant users. Our results suggest that a relatively wide dynamic range is needed for a high level of vowel recognition and a relatively small dynamic range is sufficient to maintain consonant recognition. We infer from this outcome that, if other factors were held equal, an implant patient with a small dynamic range could achieve moderately high scores on tests of consonant recognition but poor performance on vowel recognition, and that it is more likely for an implant patient with a large dynamic range to obtain high scores on vowel recognition than for an implant patient with a small dynamic range.

INTRODUCTION

In a series of papers, we have described the effects on speech understanding of processing speech signals in the fashion of a cochlear implant signal processor (Dorman et al., 1998a; Dorman et al., 1998b; Loizou et al., in press). In these experiments signals were filtered into 2-20 frequency bands, and the envelopes of the signal in each band were then extracted by full-wave rectification and low-pass filtering at 400 Hz. Sinusoids were generated with amplitudes equal to the root-mean-square (rms) energy of the envelopes (computed every 4 msecs) and frequencies equal to the center frequencies of the bandpass filters. The sinusoids were combined and output to the listeners. The results of listening tests with normal hearing listeners indicated that sentences and consonants could be recognized with 90% accuracy with as few as 5 sine wave components and that vowels from multiple talkers could be recognized with 90% accuracy with 8 sine wave components. Similar levels of identification accuracy have been achieved by patients fit with cochlear implants (Dorman and Loizou, 1998; Dorman et al., 1998a).

In the work with normal-hearing listeners described above, stimulation was at a limited number of fixed frequency locations in the cochlea. This is also the case for the stimulation provided by the three commercial cochlear-implant systems available in the United States. A fundamental issue arising from stimulation of this sort is how listeners extract the varying formant frequencies of signals when stimulation is at a small number of unvarying frequency locations along the cochlea. In previous papers we have suggested that listeners use differences in signal levels across channels to encode the location of formant frequencies (Dorman et al., 1997; Loizou et al., 1998). Our hypothesis followed an analysis of the formant frequencies of vowels from men, women, girls and boys and the corresponding channel outputs of a 6-channel signal processor (Loizou et al., 1998). A portion of those analyses are shown in Figure 1 where the spectra of the vowels /i ɪ ɛ æ/ are shown at the top of the figure and the signal levels in each channel of a 6-channel signal processor are shown at the bottom. Note how different formant frequencies are specified by differences in signal level in adjacent frequency channels. In this series of vowels, the frequency of F1 increases steadily from 254 Hz for /i/ to 703 Hz for /æ/. The progressive increase in frequency of F1 is mirrored by a change in the relative signal levels in channels 1 and 2 of the processor. The low frequency F1 of /i/ is specified by a large amplitude in channel 1 and a small amplitude in channel 2. The higher frequency F1 of /ɪ/ is specified by a smaller difference in channel 1 and channel 2 amplitudes. For /ɛ/ channel 2 has more energy than channel 1. And for /æ/, the vowel with the highest F1 in this series, the difference between the amplitudes of channel 2 and channel 1 is larger than for /ɛ/. Thus, by computing the difference in signal levels in adjacent channels a listener could encode varying formant frequencies with stimulation at unvarying cochlear locations.

On our view, an accurate encoding of channel amplitude is essential for an accurate encoding of formant frequency by devices with a small number of output channels. For patients with a cochlear implant an accurate encoding of channel amplitude may be compromised by several factors. One factor is a reduced dynamic range – the electrical dynamic range for patients may range from 3 – 4 dB in the worst cases to 20 to 30 dB in the better cases. Another factor is the number of steps available to a patient to encode intensity -- the number of steps may vary from as few as 7 to as many as 45 (Nelson et al., 1996). In an implant patient these, and many other, factors coexist. This circumstance makes it very difficult to sort out the effects of each of the factors on speech recognition. For that reason we have turned to normal-hearing listeners to assess the effects of one of the factors, reduced dynamic range, on speech recognition.

In the present experiment, signals were processed in a manner similar to a 6-channel, cochlear-implant signal processor (Zierhofer et al., 1994) and output as a sum of six sine waves with frequencies equal to the center frequencies of the analysis filters. The amplitudes of the sine waves were then compressed in a systematic fashion to produce a reduced dynamic range for stimulation. In the baseline condition, the amplitudes of the sine waves were left uncompressed. In the other conditions, the amplitudes of the sine waves were compressed to fit within a 6, 12, 18 and 24 dB amplitude range. The aim of this experiment was to assess the effect of reduced dynamic range on speech understanding.

METHOD

Subjects

The subjects were 10 graduate and undergraduate students at Arizona State University who were paid for their participation. All of the subjects had 6-10 hours of experience listening to speech reconstructed as a sum of sine waves before participating in this experiment.

Stimuli

The test material included sentences, consonants in /aCa/ environment and vowels in /hVd/ environment. The sentence material was from the H.I.N.T. test (Nilsson et al., 1994) presented without competing noise. A different set of ten sentences was used for each compression condition. The consonant test was a subset of the Iowa consonant test (Tyler et al., 1987), that is, 16 consonants in /aCa/ environment spoken by a single male speaker. Five repetitions of each consonant were used in the blocked and randomized test sequence. The vowel material consisted of the vowels in the words: “heed, hid, hayed, head, had, hod, hud, hood, hoed, who’d, heard”. Each word was produced by three men, three women and three girls. The stimuli were drawn from a set used by Hillenbrand et al. (1995). All the stimuli were presented in a completely randomized test sequence.

Envelope amplitude compression

The envelope amplitudes were computed in a manner similar to a 6-channel CIS processor and then linearly compressed to fit within a certain dynamic range. The input envelope amplitude dynamic range of each channel was mapped to a small dynamic range (6 - 24 dB) using the input-output functions shown in Figure 2. The input envelope amplitude range (Xmin-Xmax in Figure 2) was determined using statistical measures of envelope amplitudes of each channel. The maximum envelope amplitude mapped (i.e., Xmax, the saturation point of the mapping function) was determined by computing histograms of envelope amplitudes using 100 H.I.N.T sentences [the HINT sentences were scaled so that they have the same peak amplitude] as input, and choosing Xmax to include 99% of all amplitude counts of each channel. The minimum envelope amplitude mapped, i.e., Xmin, was set to the noise floor value. The input amplitude range Xmin-Xmax (which in some channels exceeded 60 dB) was mapped to the output dynamic range Ymin-Ymax (6 - 24 dB) using the input-output functions shown in Figure 2.

The envelope amplitude mapping was performed as follows. Let A_i be the envelope amplitude of channel *i (i=1,2,¼,6)*, then the compressed amplitude B_i was estimated using the equation[1]:

$$B_i = cA_i + d \tag{1}$$

where ‘c’ and ‘d’ are constants chosen such that the compressed amplitudes fall within a certain dynamic range, namely 6 dB, 12 dB, 18 dB or 24 dB. The constants ‘c’ and ‘d’ were computed according to the following equations:

$$c = \frac{Y\max - Y\min}{X\max - X\min} \tag{2}$$

$$d = Y\min - c\ X\min \tag{3}$$

where Ymax and Ymin are the maximum and minimum output envelope amplitudes mapped respectively (Figure 2). In cochlear implant processors, Ymax corresponds to the uncomfortable level and Ymin corresponds to the threshold level determined using psychoacoustic measures. For the purpose of these experiments we fixed Ymin to Xmin, i.e., Ymin=Xmin, and varied the value of Ymax so that the compressed amplitudes fell within a certain dynamic range. So, for example, the Ymax value for the 24 dB dynamic range condition was computed as $Y_{\max} = Y_{\min}\ 10^{24/20}$. In general, Ymax was computed as

$$Y_{\max} = Y_{\min}\ 10^{DR/20} \tag{4}$$

where DR is the desired dynamic range in dB (the dynamic range in dB is defined as $20\log_{10}(\text{Ymax} / \text{Ymin})$). In summary, the envelope amplitudes were compressed to a given dynamic range DR, by first computing Ymax according to Equation (4), and then incorporating the computed Ymax value in equations (2)-(3) to estimate the constants ‘c’ and ‘d’, which were in turn used in Equation (1).

It is very important to note that the above algorithm compressed all the stimuli to a *maximum* dynamic range of 6, 12, 18 and 24 dB. Statistical estimates of the dynamic range distributions, however, revealed that only a small proportion of the channel amplitudes had a dynamic range close to the maximum. Figure 3 shows, as an example, the histograms of the dynamic range of channel amplitudes computed using 20 H.I.N.T. sentences. These histograms were estimated by processing speech through 6 channels, computing the channel amplitudes using envelope detection, compressing the channel amplitudes (as per Equation 1), and estimating the dynamic range across the 6 compressed amplitudes in each 4-ms frame. As shown in Figure 3, the effective dynamic range is, on the average, much smaller than the maximum dynamic range, especially in the small dynamic range cases (6 and 12 dB). The median dynamic ranges were 2.1 dB for the 6-dB condition, 3.8 dB for the 12-dB condition, 6.18 dB for the 18-dB condition, 9.14 dB for the 24-dB condition, and 21.72 dB for the uncompressed condition. The shape of the dynamic range distribution is very important, particularly when the maximum dynamic range is very small, and is largely dependent on the mapping function. In this study, a linear mapping function (Equation 1) was used and the resulting distribution was asymmetric with a mean close to the smaller dynamic range values. A different dynamic range distribution is expected with cochlear implants, since a logarithmic-type mapping function is used.

The effect of the compression on the channel output levels is illustrated in Figure 4, which shows the channel output levels of the vowel /ε/ when compressed into a 24, 18, 12 and 6 dB dynamic range. As the dynamic range is reduced, the shape of the channel output spectrum is preserved but the peak-to-trough amplitude differences are progressively reduced. Therefore, the overall effect of the compression is a reduction of spectral contrast among channel amplitudes.

Signal Processing

The stimuli were processed as follows. The signal was first pre-emphasized (low pass below 1200 Hz with 6 dB/octave roll-off), and then bandpassed into six logarithmic frequency bands using sixth-order Butterworth filters. The center frequencies of the six bands were at 393 Hz, 639 Hz, 1037 Hz, 1685 Hz, 2736 Hz and 4443 Hz respectively. The envelope of the signal was extracted by full-wave rectification and low-pass filtering (second-order Butterworth) with a 400 Hz cutoff frequency. The channel amplitudes were estimated from the rms energy of the envelopes (computed every 4 msecs). These amplitudes were then linearly compressed according to Equation (1) to fit within a certain dynamic range, namely 6 dB, 12 dB, 18 dB or 24 dB. Sine waves were generated with amplitudes equal to the compressed amplitudes, and frequencies equal to the center frequencies of the bandpass filters. The sinusoids from each band were finally summed and presented to the listeners at a comfortable level.

Procedure

All the compressed stimuli (at 6, 12, 18 and 24 dB) were completely randomized with the uncompressed stimuli. The tests were run in the order: sentences, consonants and vowels. This order was chosen to maximize the experience of the subjects before they were tested with the difficult task of vowel identification. Practice for the sentence material consisted of the presentation of one list of ten sentences with concurrent visual display of the material for each compression condition. Practice for the vowel and consonant tests consisted of two repetitions of the stimuli with visual indication of item identity. In addition, the subjects were presented one randomized set of stimuli with feedback of correct answers. Practice was given for each compression condition.

The subjects were tested in a sound attenuated room. Responses were collected with custom software using a computer display of response alternatives and a mouse as a response key. The subjects were allowed to use a "repeat" key as many times as they wished except for the sentence tests in which only one presentation was allowed. All of the test materials were stored on computer disk and were output via custom software through a 16 bit D/A converter.

RESULTS

The results are shown in Figure 5 a-c. A repeated measures analysis of variance for the sentence material indicated a significant effect for compression ($F[4,36]=46.41$, $p<0.0001$). Post hoc tests according to Scheffe (alpha=0.05) indicated that the mean score of the normal condition (98.9%) differed from the mean score of all other conditions. The mean scores for the 6 dB and 12 dB conditions (66% and 69% respectively) did not differ.

A repeated measures analysis of variance for the Iowa consonants indicated a significant effect for compression ($F[4,36]=12.80$, $p<0.0001$). Post hoc tests according to Scheffe (alpha=0.05) indicated that the mean scores of the normal and the 24 dB condition (80.7% and 78.9% respectively) did not differ. Both scores were significantly different from the 18 dB, 12 dB and 6 dB conditions.

The results for feature analysis for the consonants are shown in Figure 6. For the feature "place of articulation" a repeated measures analysis of variance indicated a main effect for compression ($F[4,36]=32.51$, $p<0.0001$). Post-hoc tests indicated a significant difference between the percent information received in the 6 and 12 dB conditions, the 12 and 18 dB conditions, and the 24 dB and uncompressed conditions. There was no significant difference between the percent information received in the 18 and 24 dB conditions. For the feature "voicing" a repeated measures analysis of variance indicated a main effect for compression ($F[4,36]=3.99$, $p=0.009$). Post-hoc tests indicated a significant difference between the percent information received for the 6 and 12 dB conditions. There was no significant difference between the information received for the 12 dB conditions and the other conditions. For the feature "manner" repeated measures analysis of variance indicated no main effect of compression. The mean information received was 99.5% in all conditions.

A repeated measures analysis of variance for the multi-talker vowel test indicated a significant effect for compression ($F[4,36]=63.78$, $p <0.0001$). Post hoc tests according to Scheffe (alpha=0.05) indicated that the mean score for the normal (uncompressed) condition differed from all other conditions. The mean scores for the 18 dB and 24 dB conditions (63.8% and 68.4% respectively) did not differ.

DISCUSSION

There was a significant effect of compression for all test materials. The magnitude of the effect differed as a function of the test material. Sentence recognition was extremely robust even for the 6-dB condition where the median dynamic range was only 2 dB. The scores for vowel recognition and recognition of the consonant feature "place of articulation" were affected the most. Both vowel recognition and the recognition of consonant place of articulation require a listener to resolve the locations of spectral peaks in the input signal. It is reasonable to suppose that the mechanism underlying the poor recognition of vowels and consonant place of articulation is a loss of spectral contrast following the reduction in signal dynamic range. If this is the case, then this outcome fits with our view that, when speech is processed into a small number of fixed-frequency channels, listeners must use across-channel differences in signal level to code differences in signal frequency.

The robust recognition of consonant manner and voicing in the face of a large reduction in signal dynamic range can be rationalized on the view that one set of the acoustic cues which specify manner and voicing are found in the gross shape of the amplitude envelope. These envelope shapes are found concurrently in several channels of a multi-channel processor and, as a consequence, are relatively robust in the face of reductions in signal resolution. Reception of voicing and manner information is robust even in single-channel processors in the absence of spectral information (e.g., Rabinowitz and Eddington, 1995) and in single-channel syllabic compression processing (e.g., Souza and Turner, 1996). It is of interest to note that consonant voicing is normally signaled by a combination of a temporal, or envelope, feature and the presence or absence of noise excitation. However, in our stimuli there was no noise excitation for voiceless signals. Therefore, voicing was signaled principally by the shape of the amplitude envelope.

The effects reported above bear, at least in a qualitative manner, on the speech perception skills of patients fit with a cochlear implant. Most generally, patients with a very small dynamic range should perform worse than patients with a very large dynamic range on tests of speech recognition. Indeed the effects of a reduced dynamic range may be exaggerated for patients with implants. In our experiment, the signals were presented at a moderate listening level and even the lowest level signals were well above threshold. For an implant patient low-level signals are near threshold because signals are mapped between threshold and uncomfortable loudness. For this reason the effects of a relatively small dynamic range may be larger for patients than for the normal-hearing listeners tested in our experiment. Whether this speculation is correct or not, the outcomes for vowels, consonants and sentences reported in this paper suggest that dynamic range should affect the speech perception abilities of cochlear implant users. This appears to be the case.

Blamey et al. (1992) reported a small but significant correlation (r =0.398) between dynamic range and speech understanding for patients fit with cochlear implants. It is likely that the very modest value of the correlation coefficient is due to factors other than dynamic range which interact with dynamic range to determine the performance of a given patient. Thus, a patient with a large dynamic range who, other factors equal, should perform relatively well on tests of speech understanding could have a relatively short depth of electrode insertion and could have only a relatively small number of discriminable intensity steps within his/her dynamic range. The latter two factors, as well as many others, would act to limit speech understanding. Thus, it is not surprising to find only a very modest correlation between dynamic range, per se, and speech understanding for patients fit with a cochlear implant.

Implant patients differ from normal-hearing listeners not only in terms of reduced dynamic range but also in terms of the number of discriminable intensity steps within that range. Nelson et al. (1996) have described patients who have as few as 7 intensity steps across their dynamic range. Other patients have as many as 45 intensity steps across their dynamic range. The number of steps available to code differences in intensity will, no doubt, be a factor, in addition to the dynamic range, which can affect implant performance. It is possible, for instance, that the number of steps available in a patient's dynamic range may be a more important factor for speech perception than the dynamic range itself. When looking for factors which affect implant performance, both dynamic range and the number of discriminable intensity steps within that range should be taken into account. In the present study, we only investigated the effect of reduced dynamic range on speech perception, thereby assuming that there were a large number (infinite) of steps within the given dynamic range. Further studies are needed to investigate the effect of the interaction between dynamic range and number of steps on speech understanding.

Acknowledgment

The authors would like to thank Fan-Gang Zeng, William Rabinowitz and Peter Blamey for providing valuable suggestions that helped improve the manuscript. This research was supported by Grants No. R01 DC03421 and R01 DC000654-6 from the National Institute on Deafness and other Communicative Disorders, NIH.

Footnotes

[1] Equation (1) describes linear compression. Logarithmic compression is typically used in cochlear implant processors in order to match the loudness between the acoustic and electrical stimulation (Zeng and Shannon, 1992). Our simulation of the compression function does not represent accurately the loudness growth function found in electrical stimulation. The consequence of that is that sounds, which would be near threshold for an implant user, are much louder for a normal-hearing listener.

REFERENCES

Blamey, P., Pyman, B., Gordon, M, Clark, G., Brown, A., Dowell, R. and Hollow, R. (**1992**). "Factors predicting postoperative sentence scores in postlinguistically deaf adult cochlear implant patients," *Ann. Otol. Rhinol. Laryngol*. **101**, 342-348.

Dorman, M, Loizou, P. and Rainey, D. (1997). "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**, 2403-2411.

Dorman, M., Smith, L. and Parkin, J. (**1993**). "Loudness balance between acoustic and electric stimulation by a patient with a multichannel cochlear implant." *Ear Hear.,* **14** (4), 290-292.

Dorman, M. and Loizou, P. (1998). "The identification of consonants and vowels by cochlear implant patients using a 6-channel CIS processor and by normal-hearing subjects using simulations of processors with 2-9 channels," *Ear Hear.* **19(2),** 162-166.

Dorman, M., Loizou, P. and Fitzke, J. (1998a). "The identification of speech in noise by cochlear implant patients and normal hearing listeners using 6-channel signal processors," *Ear Hear.* **19(6),** 481-484.

Dorman, M., Loizou, P., Fitzke, J. and Tu, Z. (1998b). "The recognition of sentences in noise by normal-hearing listeners using simulations of cochlear-implant signal processors with 6-20 channels, "*J. Acoust. Soc. Am*., **104(6),** 3583-3585.

Hillenbrand, J., Getty, L., Clark, M. and Wheeler, K. (**1995**). "Acoustic characteristics of American English vowels," *J. Acoust. Soc. Am*. **97**, 3099-3111.

Loizou, P., Dorman, M. and Powell, V. (1998). "The recognition of vowels produced by men, women, boys and girls by cochlear implant patients using a six-channel CIS processor," *J. Acoust. Soc. Am.* **103**(2), 1141-1149.

Loizou, P., Dorman, M. and Tu, Z. (in press). "On the number of channels needed to understand speech," *J. Acoust. Soc. Am*.

Nelson, D., Schmitz, J., Donaldson, G., Viemeister, N. and Javel, E. (**1996**). "Intensity discrimination as a function of stimulus level with electric stimulation." *J. Acoust. Soc. Am*. **100** (4), 2393-2414.

Nilsson, M., Soli, S. and Sullivan, J. (**1994**). "Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise, " *J. Acoust. Soc. Am.* **95**, 1085-1099.

Rabinowitz, W. and Eddington, D. (**1995**). "Effects of channel-to-electrode mappings on speech reception with the Ineraid cochlear implant," *Ear and Hearing*, **16**(5), 450-458.

Souza, P. and Turner, C. (**1996**). "Effect of single-channel compression on temporal speech information," *Journal of Speech and Hearing Research*, **39**(5), 901-911.

Tyler, R., Preece, J. and Lowder, M. (**1987**). The Iowa audiovisual speech perception laser videodisc. *Laser Videodisc and Laboratory Report*, Dept. of Otolaryngology, Head and Neck Surgery, University of Iowa Hospital and Clinics, Iowa City.

Wilson, B., Finley, C., Lawson, D., Wolford, R., Eddington, D and Rabinowitz, W. (**1991**). “Better speech recognition with cochlear implants,” *Nature* **352**, 236-238.

Zeng, F.G. and Shannon, B. (**1992**). “Loudness balance between electric and acoustic stimulation,” *Hear. Research*, **60**, 231-235.

Zierhofer, C., Peter, O., Bril, S., Pohl, P., Hochmair-Desoyer, I., and Hochmair, E. (**1994**). “A multichannel cochlear implant system for high-rate pulsatile stimulation strategies,” in I.Hochmair Desoyer and E. Hochmair (Eds), *Advances in Cochlear Implants*. Vienna: International Interscience Seminars, 204-207.

Figure Captions

Figure 1. (top) Spectra of the vowels in "beet, bit, bet, bat" derived from a 20-ms window centered at the midpoint of the vowel, and 14 pole LPC analysis. (bottom) Output levels of a 6-channel Continuous Interleaved Sampling (CIS) processor for the vowels in "beet, bit, bet, bat". Sequential points along each function represent the output levels of channels 1-6.

Figure 2. Input-output compression functions used in this study. The thick line shows the input-output function corresponding to the uncompressed condition. In this condition, Ymax=Xmax, and Ymin=Xmin. The four thin lines show the input-output functions used for the 24, 18, 12, and 6 dB dynamic range conditions. These functions were obtained by fixing Ymin and varying the value of Ymax so that the amplitudes fell within a certain dynamic range. For example, the 24 dB input-output function was obtained by setting Y_{max1} equal to $Y_{max1} = Y_{min}\, 10^{24/20}$. The x- and y-axes are shown in linear scale.

Figure 3. Histograms of the dynamic range of channel amplitudes computed using 8,300 frames of speech (20 H.I.N.T. sentences) and compressed to the indicated dynamic ranges. The abscissa shows the dynamic range computed across the six rms amplitude levels of each frame. The histograms were estimated by bandpass filtering speech into 6 frequency bands, computing the channel amplitudes in each band through envelope detection, compressing the channel amplitudes (as per Equation 1) into the indicated dynamic ranges (6-24 dB), and estimating the dynamic range across the 6 compressed amplitudes in each 4-ms frame. In the "Full range" case, the channel amplitudes were not compressed.

Figure 4. The channel output levels of the vowel /ε/ ("head") when compressed into a 24, 18, 12 and 6 dB dynamic range. The top left panel shows the envelope amplitude levels before compression. As it can be seen, the shape of the channel output spectrum is preserved, however the peak-to-trough amplitude differences are progressively reduced as the dynamic range is reduced.

Figure 5. Mean percent correct scores on vowel, consonant and sentence recognition as a function of dynamic range. “Full range” corresponds to the case in which the amplitudes were left uncompressed. Error bars indicate standard deviations.

Figure 6. Percent information received for the features “place”, “manner” and “voicing” as a function of dynamic range. “Full range” corresponds to the case in which the amplitudes were left uncompressed. Error bars indicate standard deviations.

FIGURE 1

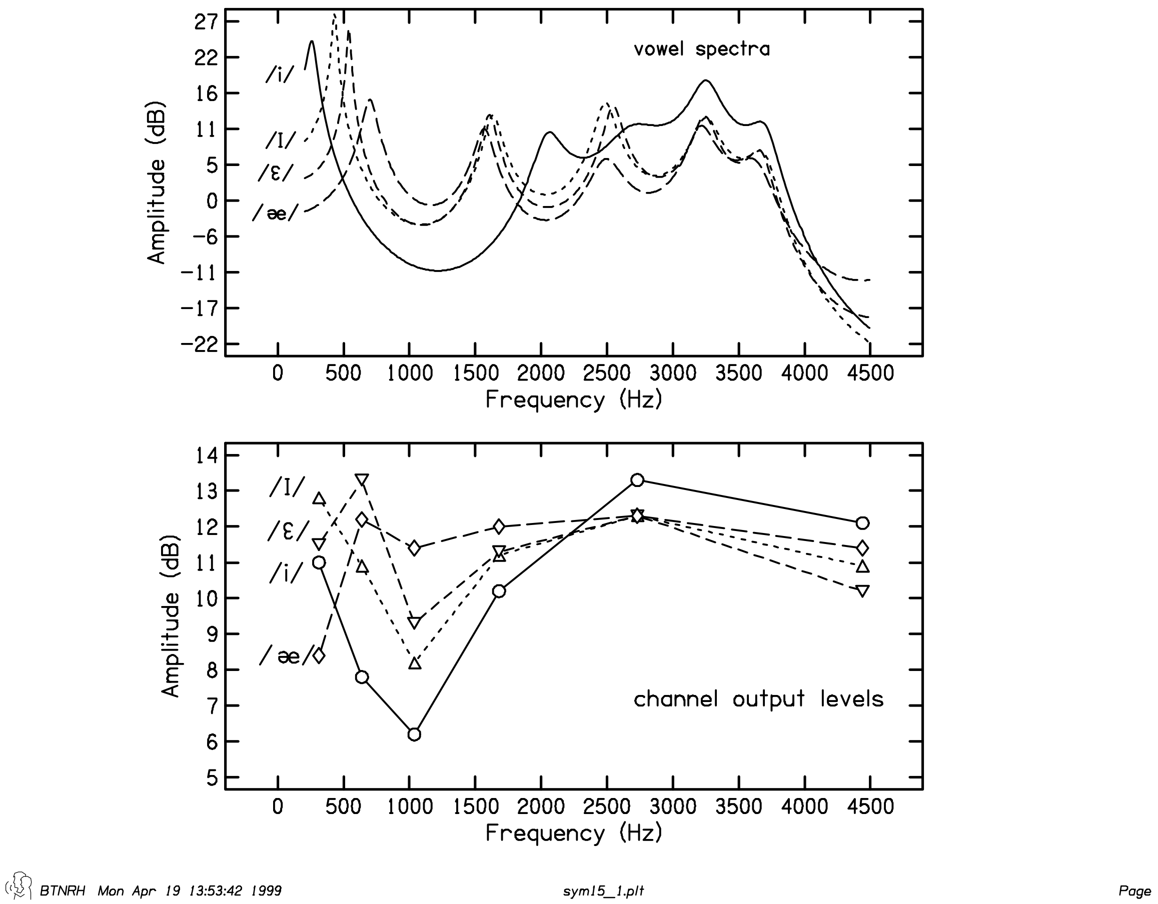
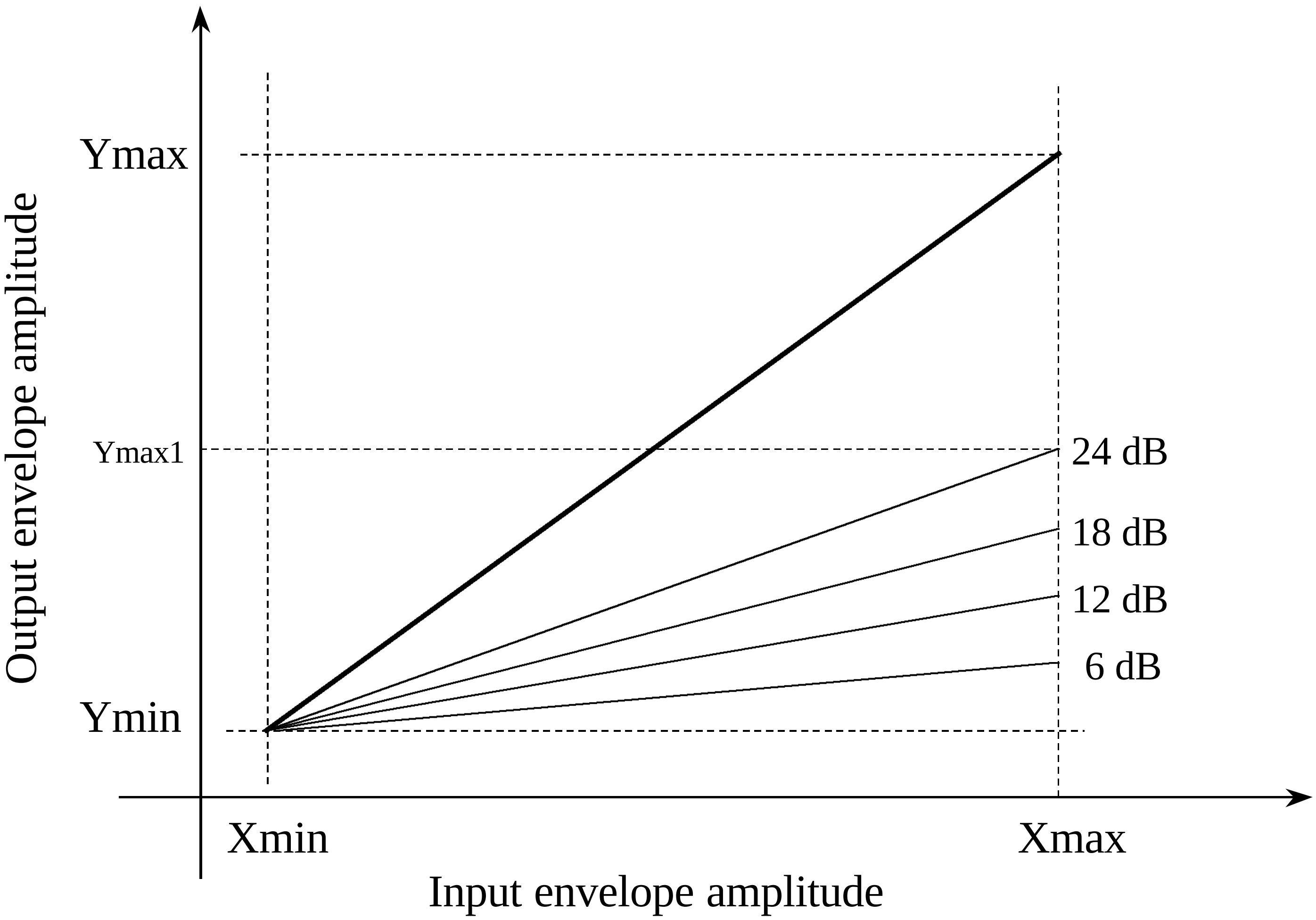

vowel spectra
/i/
/I/
/ɛ/
/æ/
Amplitude (dB)
Frequency (Hz)
channel output levels
BTNRH Mon Apr 19 13:53:42 1999
sym15_1.plt
Page 1

FIGURE 2


Ymax
Ymax1
Ymin
Output envelope amplitude
24 dB
18 dB
12 dB
6 dB
Xmin
Xmax
Input envelope amplitude

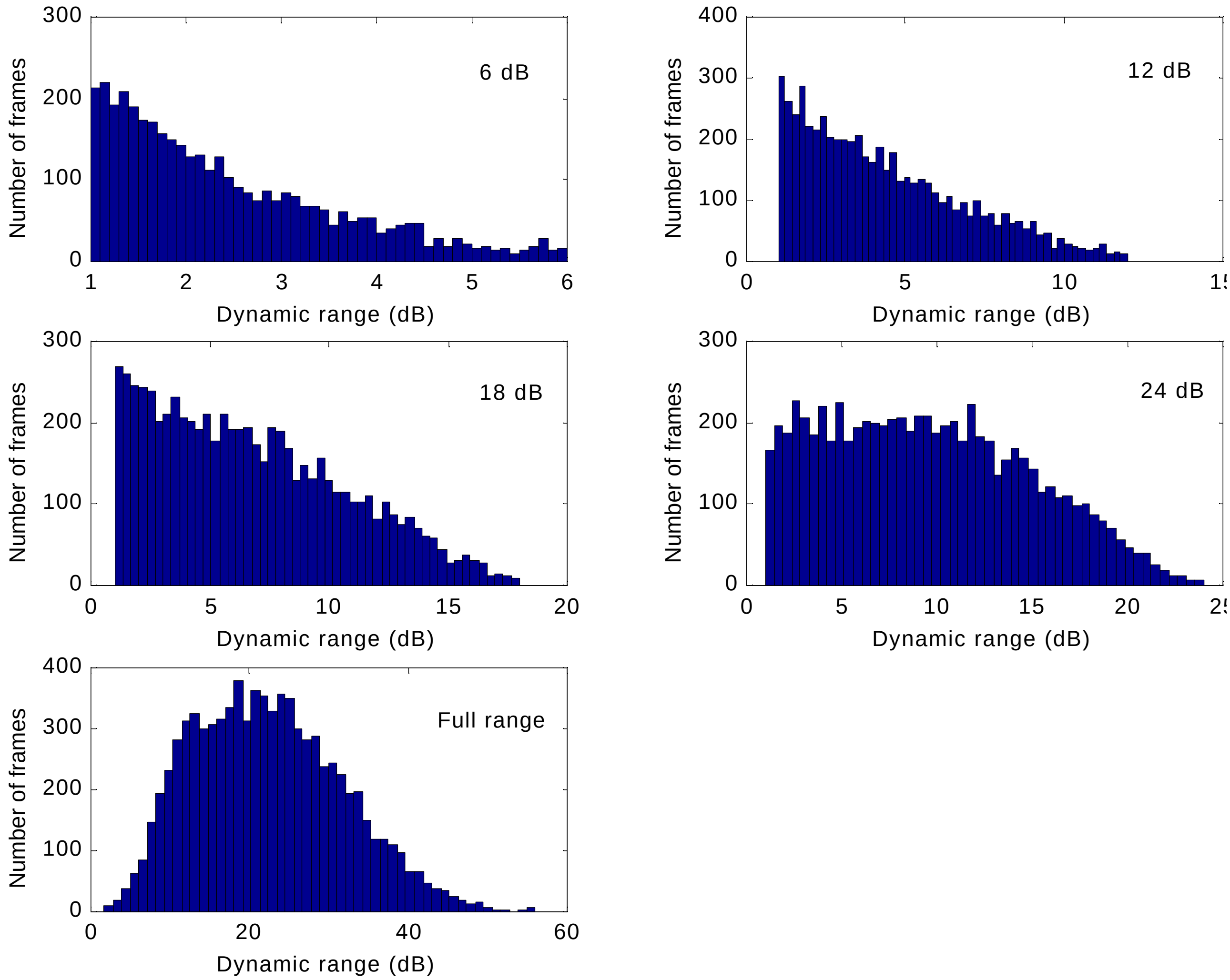
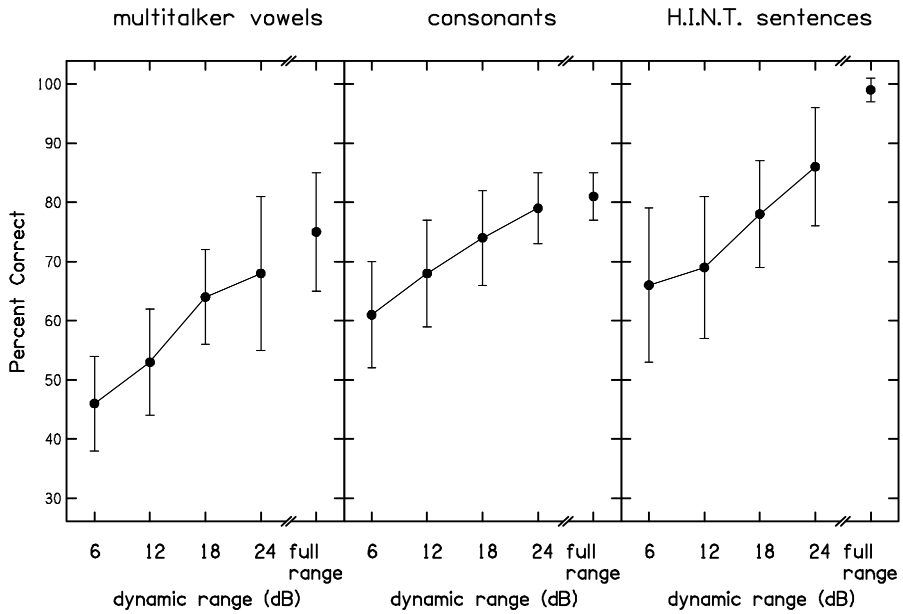

6 dB
12 dB
18 dB
24 dB
Full range
Number of frames
Dynamic range (dB)

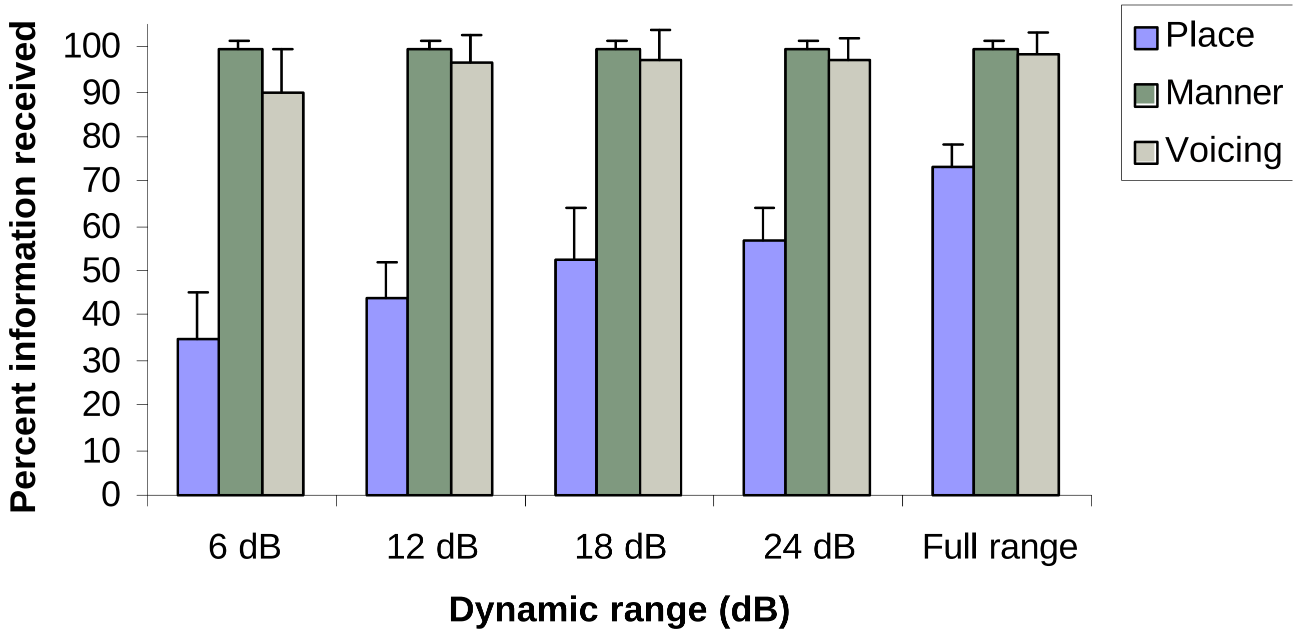
FIGURE 4

FIGURE 5



multitalker vowels
consonants
H.I.N.T. sentences
Percent Correct
100
90
80
70
60
50
40
30
6
12
18
24
full range
dynamic range (dB)

FIGURE 6



Percent information received
100
90
80
70
60
50
40
30
20
10
0
6 dB
12 dB
18 dB
24 dB
Full range
Dynamic range (dB)
Place
Manner
Voicing