Department of Veterans Affairs

Journal of Rehabilitation Research and Development Vol. 30 No. 1, 1993
Pages 117–128

Effects of noise and noise reduction processing on the operation of the Nucleus-22 cochlear implant processor

Mark R. Weiss, MSEE
Center for Research in Speech and Hearing Sciences, City University of New York Graduate School, New York, NY 10036

Abstract—Cochlear implants, like other types of auditory sensory aids, become increasingly ineffective with increasing ambient noise levels. One method of signal processing to reduce additive random wideband noise, the INTEL method, has been used to good effect as an input preprocessor for the Nucleus-22 cochlear implant. The implant's own signal processor estimates and encodes pitch frequency and the frequencies of Formants 1 and 2. The study reported here shows that additive noise results in substantial deviations in formant frequency estimates from those that would be observed in the absence of noise. When noisy speech is preprocessed by the INTEL method to reduce noise intensity, the deviations in the frequency estimates for Formant 2 are substantially reduced.

Key words: *cochlear implants, INTEL method, noise reduction, preprocessed noisy speech, signal processing.*

INTRODUCTION

Cochlear implants produce a perception of sound through electrical stimulation of auditory nerve fibers. Sounds received by the implant system microphone are transformed into patterns of electrical currents that are delivered either to a single electrode at the round window, or to one or more electrodes that are inserted into the cochlea itself. In general, two approaches are in use for converting audio signals into stimulus currents. In some devices the analog signals in one or more frequency bands are encoded, usually after some form of amplitude compression or frequency shaping has been imposed on them. Others encode information-bearing features of speech that have been extracted from the signals. Whichever coding method is used, when noise accompanies the received sounds the efficacy of the implant as an aid to speech perception is rapidly reduced. The loss of efficacy in noisy environments is similar to that experienced by users of amplification-type hearing aids and elicits similar complaints.

One of the most widely used multiple electrode cochlear implant systems is the Nucleus 22-channel device made by Cochlear, Pty. Ltd. (New South Wales, Australia). Its signal processor extracts intensity, voicing, pitch, and formant information from received sounds, encodes these features as pulsatile patterns, and transmits them to appropriate electrodes in the implant. The frequencies of the first two formants are estimated as the zero crossing rates for the dominant spectral peaks in two bands, which nominally are from 300 Hz to 1,000 Hz for Formant 1, and from 800 Hz to 4,000 Hz for Formant 2. Some noise immunity is provided in the two models of this implant that are in use. In the wearable speech processor (WSP III) model, stimulation is suppressed for sounds below a minimum average level. In the minispeech processor (MSP) model, the long-term average signal level is computed in each of several bands, and this constant value is subtracted

Address all correspondence and requests for reprints to: Mark R. Weiss, Center for Research in Speech and Hearing Sciences, City University of New York, 33 West 42nd Street, New York, NY 10036.

from the instantaneous signals in these bands. The band outputs are then recombined. While the frequency measurement and noise reduction methods are simple in comparison with more powerful procedures that are available, they apparently are adequate in relatively quiet environments. However, it has been reported that changes in user performance with the noise suppression system active are variable, and that users of the WSP III model tend not to activate noise reduction, while users of the MSP model vary in their use of it (1).

The effects of noise on speech perception are substantially different for sounds that are heard directly as compared with sounds received through cochlear implants. In the former case, it is primarily the masking and critical band characteristics of the ear that determine how noise will affect perception. In the latter case, noise-induced errors in the feature measurements will result in the generation of incorrect stimulation patterns, resulting in misperceptions of the encoded sounds. Such errors can be reduced by using more robust methods of extracting the selected speech features or, alternatively, by processing noisy speech so as to improve the signal-to-noise ratio. The latter approach was taken by Hochberg, et al. (2) who studied the effect of using the INTEL noise reduction method (3,4) to process noisy speech that was presented to users of the Nucleus-22 implant. Consonant-vowel-consonant (CVC) words drawn from isophonemic word lists (5) were presented to cochlear implantee subjects at signal-to-noise ratios in the range 25 dB to −10 dB within a band from 0 to 4,500 Hz. Substantial improvements in perception accuracy were obtained for signals in additive speech spectrum noise at signal-to-noise ratios down to 0 dB. Significantly, the average shift in subject performance was reasonably consistent with the estimated reduction of noise in the processed signals.

In view of these results, it was of interest to examine the effect INTEL processing of the CVC words had on the formant measurement accuracy of the Nucleus-22 processor, and in particular to identify where and to what extent formant measurement errors were reduced. The next section of this paper describes the noise reduction process and the manner in which it was applied to the test stimuli in the experiments reported by Hochberg, et al. (2). Next, the method by which formant measurement errors were determined is discussed. This is followed by a presentation and a discussion of the results of the analyses. Finally, the implications for the design of cochlear implant systems are considered.

Noise Reduction Method

The INTEL noise reduction process is a "transformation subtraction" type of procedure. In this approach to noise reduction, a function that represents the noise is subtracted from transforms of the noisy speech signals. The modified signal transforms are then reconverted to time waveforms by use of an inverse of the original transformation procedure. In systems that use this approach, the noise reference function usually is generated from transforms of the input signals that ideally contain little or no speech energy. In multiple-channel noise reduction systems, the data that are needed to generate this function are obtained by use of one or more additional microphones that receive versions of the noise in which signal components are weak or absent, that is, adaptive noise cancelers (6). In single-channel systems, such as INTEL, the noise reference function is derived from the noisy speech signal itself, and so can be defined only when speech is absent or is negligible relative to the noise. Consequently, in such systems, the noise reference function cannot represent the instantaneous variations of the noise, but only their statistical characteristics. In INTEL, the noise reference function is based solely on the average transform of the noise.

INTEL operations (see **Figure 1**) are performed on successive overlapped segments of the input signal. Four Fourier transforms are computed during the processing of each segment. The first two of these are needed to generate a function on which the noise reduction operations are performed. First, the complex spectra of the signal are obtained and converted into magnitude and phase transforms. Then, an inverse transform of the square-root of the magnitude frequency spectrum is computed, and the magnitude of that transform is obtained. For convenience, the resulting function is referred to here as the period spectrum, since the locations of peaks in this function are linearly related to the periods of periodic components in the input signal.

To generate the noise reference function, the average period spectrum of the noise is computed, segmented at some period, Tc, into two regions, and scaled in each region. The value of Tc (usually Tc $<$ 1 msec) and the scale factors used for each

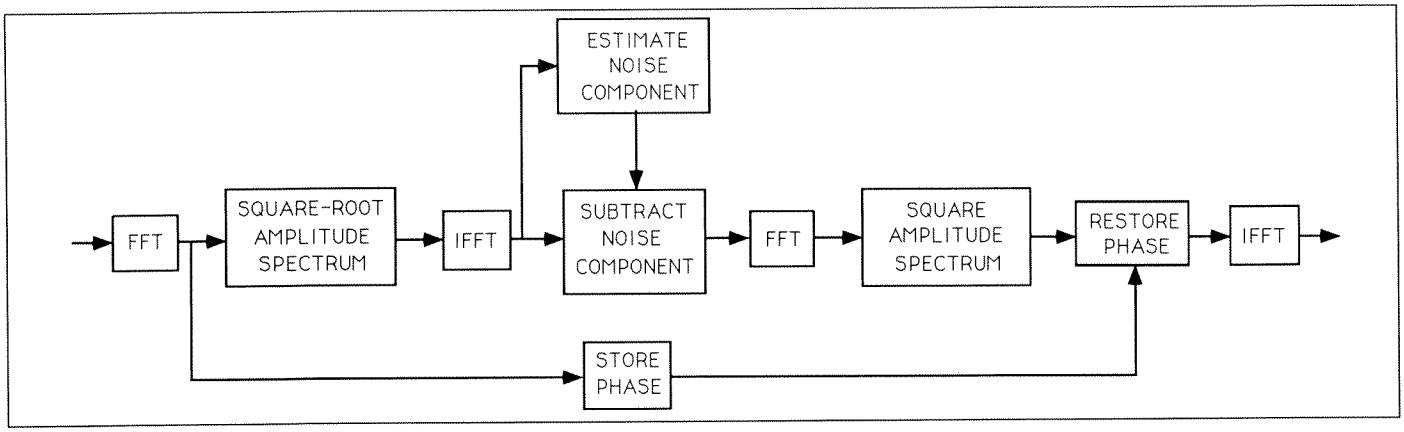


Figure 1.
INTEL method of noise reduction.

region are selected to maximize output signal quality for a given amount of noise reduction. When the noise reduction process is active, the noise reference function is subtracted from the period spectra of the input. Resulting differences that are negative are set to zero. Then the modified period spectra are converted by direct Fourier transformation to the frequency domain, squared to restore original amplitude relationships, and combined with the original phase frequency spectra to generate new complex spectra. These are transformed to the time domain to produce output signal segments which are combined by use of an overlap-add type procedure to generate a continuous output signal. The overlap-add procedure is used to reconstruct a signal or a processed version of a signal by summing the regenerated segments of the output that correspond to overlapped segments of the input (7).

The effects of the INTEL process on speech signals and on noise depend primarily on the values set for the noise-reference function scale factors. While the process reduces the level of noise, it also reduces the level of relatively weak speech sounds (e.g., weak fricatives or plosives). It also tends to weaken disproportionately the weaker components of speech sounds. These effects increase in proportion to the amount of noise reduction that is achieved. Generally, the larger the scale factors are made, the greater will be the reduction of noise and, concurrently, the greater will be the distortion in the regenerated speech. Where high levels of distortion are acceptable, as in the case of limited vocabulary speech, noise reductions of up to 14 dB can be achieved, with scale factors usually in the range 0.6 to 0.85. With scale factors in the range 0.2 to 0.4, noise can be reduced by up to 4 dB with little perceptible distortion in the processed speech sounds.

In the experiments described in Hochberg, et al. (2) the INTEL process was set to reduce noise by 7 dB on average over the range of input signal-to-noise ratios. For this setting, the period spectrum low-band cutoff was 0.2 msec, and the low-band and high-band scale factors were 0.6 and 0.65. The amount of noise reduction was estimated by measuring the noise level before and after processing when speech was absent. Since the INTEL process is nonlinear, the measured reduction in noise level does not necessarily indicate that the average output signal-to-noise ratio was increased by 7 dB. However, it is noteworthy that the test results in Hochberg, et al. (2) indicate that the average shift between performance intensity (PI) functions for the processed and unprocessed stimuli, measured at the signal-to-noise ratio which obtained 50 percent of the maximum score, was 5 dB. This result is in reasonable agreement with the shift that would be expected for a signal-to-noise ratio enhancement of 7 dB.

A typical example of the noise reduction that was obtained for these settings of the INTEL process is presented in **Figure 2**. The spectrogram of the phrase "the word is ways," at a signal-to-noise ratio of 25 dB, is shown at the top. The middle spectrogram shows the result of adding speech-spectrum shaped noise at a level 10 dB below the

Journal of Rehabilitation Research and Development Vol. 30 No. 1 1993

speech. At a signal-to-noise ratio of 10 dB, Formant 1 is partially obscured in regions where it is weak, and Formant 2 is almost completely obscured, especially during the /w/ glides. As seen in the lower spectrogram, after processing, Formant 2 is more apparent both during the transitions and in the regions of less rapid change.

METHODS

The formant frequency data required for this study were obtained by use of a Cochlear Pty., Ltd. simulator of the Nucleus-22 processor. The simulator, which ran in an IBM/AT-type personal computer, performed the signal conditioning, filtering, and formant measurement operations that are implemented in both the WSP III and the MSP model processors, and provided formant frequency measurements at intervals of 1 msec. Three sets of measurements were obtained. Two of these were at a signal-to-noise ratio of 10 dB, one set with noise reduction processing and the other without processing. Measurement errors for each test condition were determined relative to a reference set of data. Ideally, this set should consist of the true formant frequency values for the speech test material. However, the true formant values are unknown and the estimated formant frequency values obtained with the simulator in quiet were used.

The signal-to-noise ratios selected for these sets were chosen to permit the results of these analyses to be compared with data obtained in the experiments described in Hochberg, et al. (2). The signal-to-noise ratio of 25 dB used for the reference data set is the same as the maximum signal-to-noise ratio used in the cited experiments and was used to establish the upper performance limits for the tested subjects. The signal-to-noise ratio of 10 dB used for the test data sets is approximately at the middle of the test range used in the cited experiments.

The differences in the detectability of Formants 1 and 2 that are apparent in the middle and lower spectrograms of **Figure 2** are reflected in the performance of the Nucleus-22 formant frequency extractor. As can be seen in **Figure 3**, the measurements for Formant 1 at a signal-to-noise ratio of 25 dB form a reasonably smooth track for Formant 1 through the glide /w/ and the diphthong /eI/ in the test word "ways." Formant 2 values are tracked

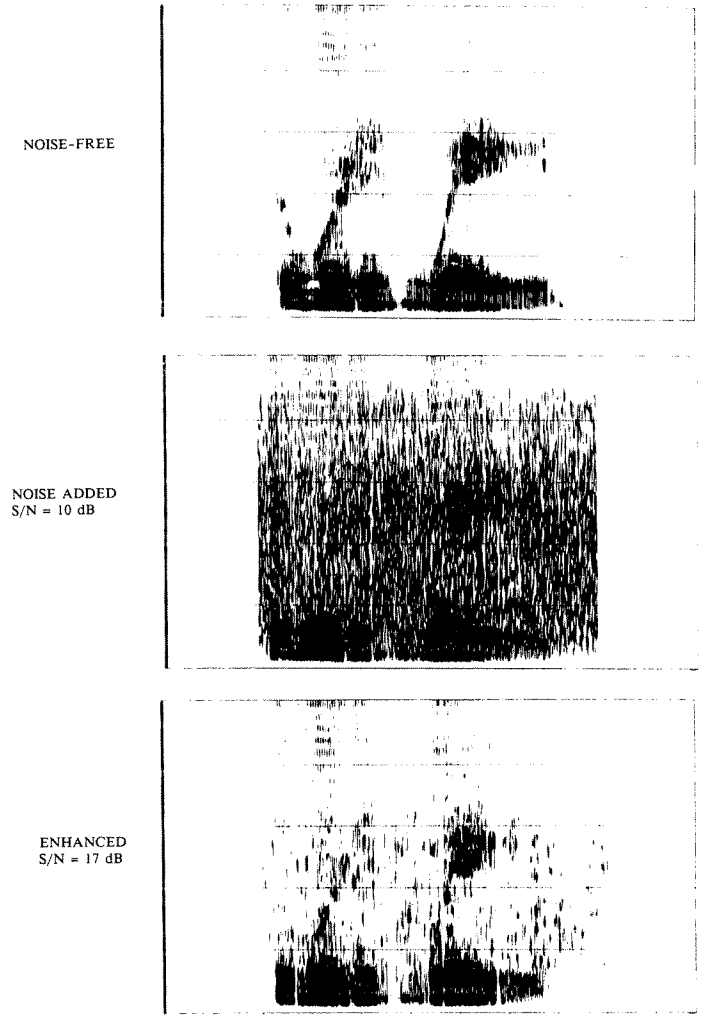


Figure 2.
Spectrograms of "The Word is Ways" before and after INTEL processing.

well through the glide, but there appears to be some uncertainty during the diphthong. At a signal-to-noise ratio of 10 dB, the tracking of Formant 1 is still reasonably good, with some variability evident at the end of the diphthong. However, the first half of the Formant 2 transition during the glide is lost. After processing, the extraction of Formant 1 is essentially the same as before processing, but now the entire Formant 2 transition has been tracked.

In the experiments described in Hochberg, et al. (2) the test stimuli were drawn from recordings of 15 isophonemic CVC word lists as spoken by a female talker, with each list composed of 10 words. Speech-spectrum-shaped noise was added at a constant level, and the speech level was adjusted to

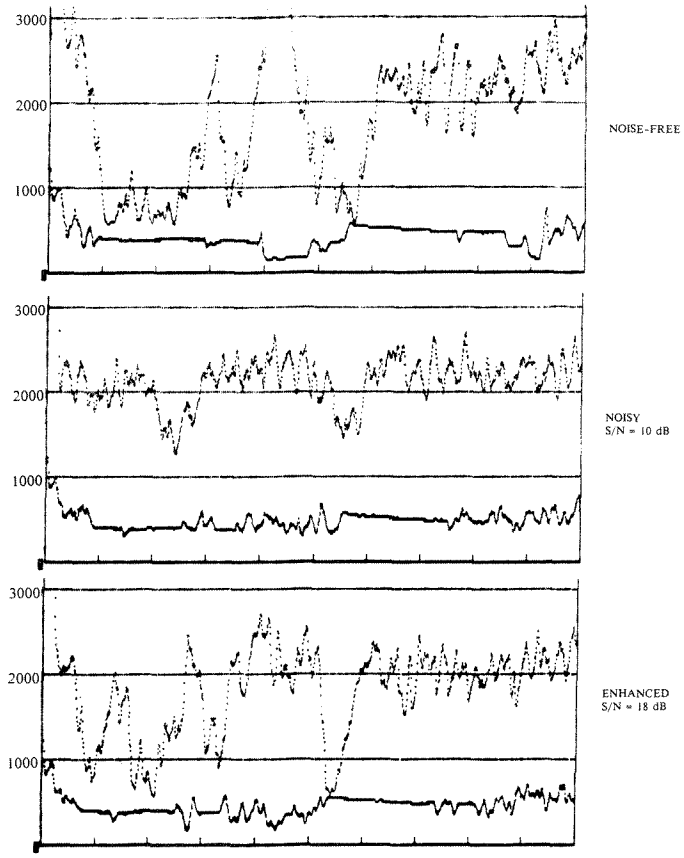


Figure 3.
Estimated frequencies of Formants 1 and 2 for "Word is Ways" before and after INTEL processing.

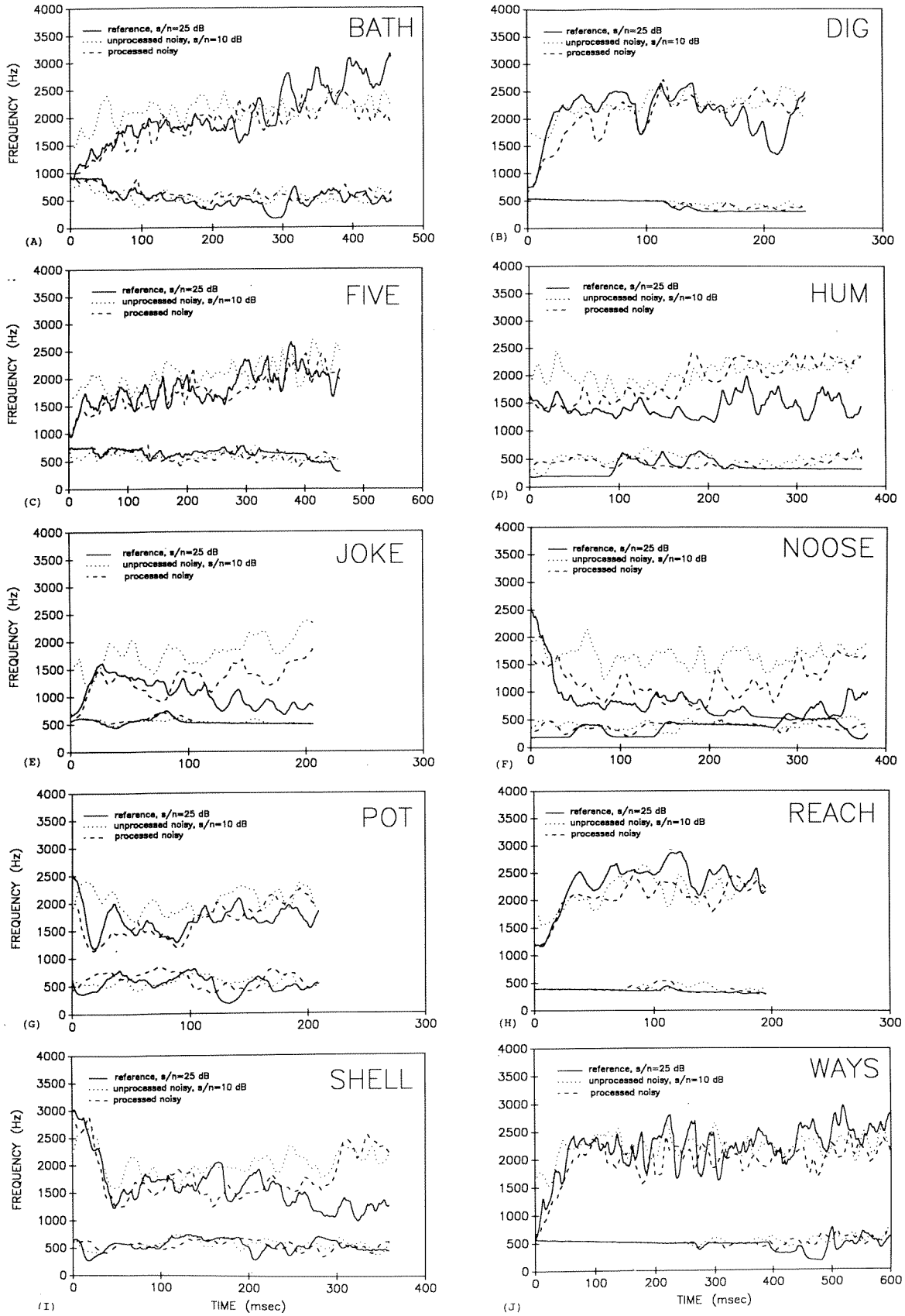
achieve the desired signal-to-noise ratio. The formant frequency data that were used in the study described here were extracted from the recording of word list No. 5, in which the test words were: *bath, dig, five, hum, joke, noose, pot, reach, shell,* and *ways.* The intensity and spectral distribution of the noise and the settings of the INTEL process parameters were the same as in the cited experiments. Inputs to the simulator were 5-kHz low-pass filtered, sampled at a rate of 10 kHz, and converted to digital form with 12-bit accuracy.

For all but one of the test words in the reference set ("hum"), the measured frequencies of one or both of the formants changed rapidly either during the transition from the initial consonant to the vowel or during the consonant itself. It was expected that at any given signal-to-noise ratio, formant measurement errors would be greater during these rapid changes than during periods when the formant frequency was changing slowly or not at all, and that this would be especially so when these changes occurred in regions where speech intensity is lower than it is during the vowel nucleus. Conversely, it was anticipated that an increase in signal-to-noise ratio would reduce errors in these regions more rapidly. To observe whether INTEL processed speech exhibited such an effect, the extracted formant tracks for each test token were divided into an initial region and a central region. The initial region was defined as beginning at the start of the transition to the central vowel and ending at the point at which the frequencies of both formants were within 10 percent of their estimated frequencies at the vowel nucleus. The central region encompassed the remainder of the vowel. The difference between the test and reference values, determined at 1 msec intervals, were used in the computation of three types of errors within each region: the average absolute error, the root mean square (rms) error, and the maximum absolute error. For the absolute error and rms errors the difference measures at each point were expressed as a percent of the reference value.

RESULTS

Formant Data

Figures 4a, 4b, 4c, 4d, 4e, 4f, 4g, 4h, 4i, and **4j** present graphically the formant measurements made by the Nucleus-22 simulator for each of the test words under each of the study conditions. It is apparent that the reference data for Formant 2 exhibit considerable variability, both during the onset of the voiced speech sounds and during the relatively slowly changing segments of the central vowels. These variations indicate an intrinsic weakness in the method of estimating formant frequencies by counting zero crossings, particularly as applied to extracting vowel formants from speech with a high fundamental frequency. Comparisons of the reference condition measurements with formant frequency estimates obtained from spectrograms and from linear predictive coding (LPC) analyses also show occasional gross errors in average values for one or both formants. However, since these data were used to represent the normal "noise free" performance of the formant extractors, then for the purposes of these analyses the reference data were considered to be error free. Thus, an initial region



Figures 4a–4j.
Estimated frequencies of Formants 1 and 2 for 10 test words before and after INTEL processing.

was determined in accordance with the definition stated earlier even when the initial rapid variations in formant frequency did not correspond to events that could be observed in a spectrogram.

The formant tracks for the test conditions exhibit short-term variations similar to those seen for the reference condition. The general trends of the test and reference condition measurements are more evident in **Figures 5a, 5b, 5c, 5d, 5e, 5f, 5g, 5h, 5i,** and **5j**, which present the formant data fitted with tenth-order polynomials. During the first halves of the test, words the smoothed Formant 2 tracks for the test conditions tend to parallel the reference tracks, with the tracks for the processed data usually closer to the reference tracks than are the tracks for the unprocessed data. The deviation between the test condition tracks and the reference tracks tends to increase during the second halves of the words.

Error Analyses

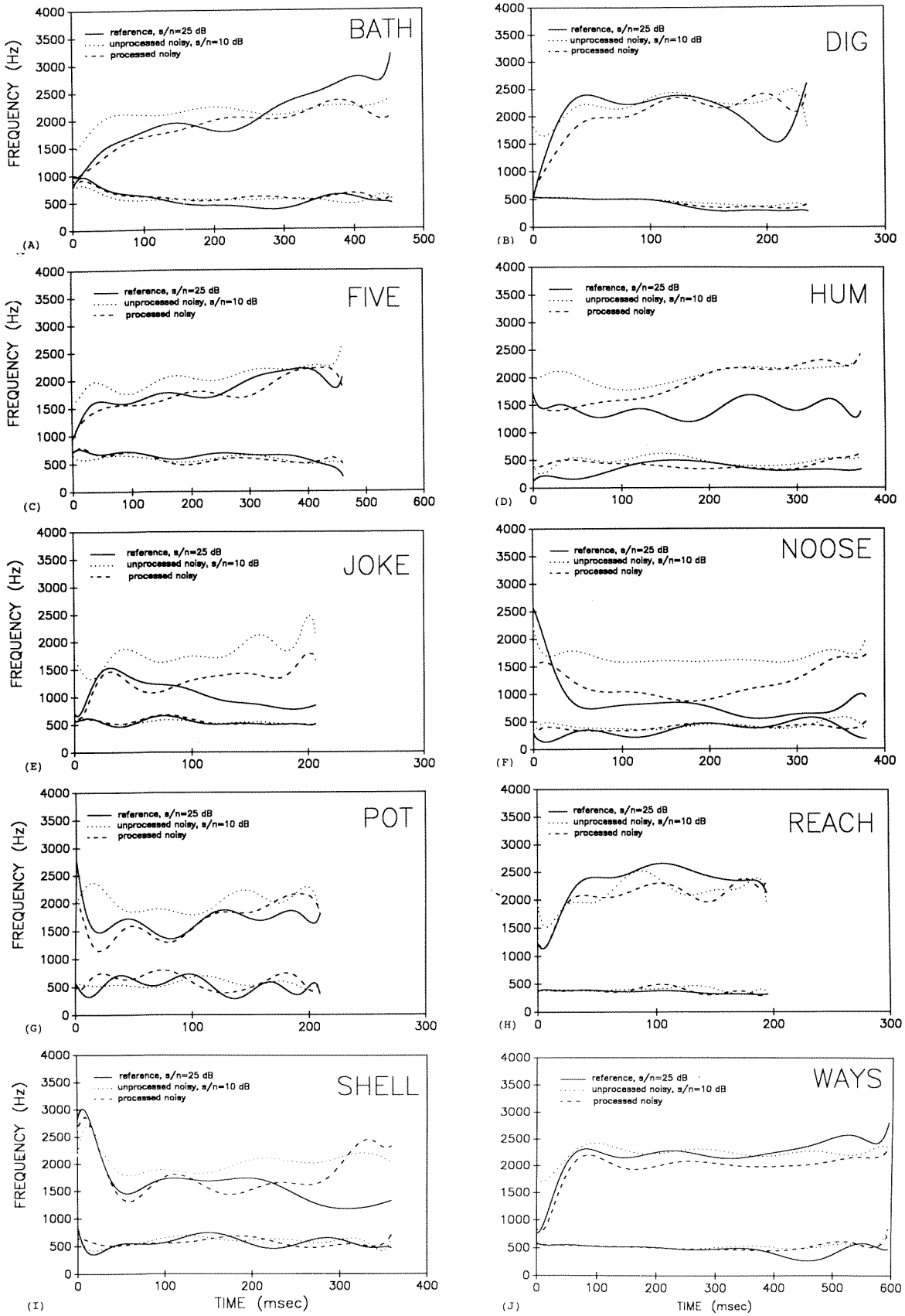
The results of the error analyses are presented in **Table 1, Table 2,** and **Table 3**. Reduced error rates in the formant measurements made on the processed sounds are evident for both formants and for all error measures. Analyses of these data, using Wilcoxon matched-pairs signed-ranks statistics, show that the effect of noise reduction processing on Formant 1 was not significant ($p > 0.15$) in any of the error measures in either the initial or central region. On the other hand, the effect of processing is seen to be highly significant for Formant 2 in the initial region in all three error measures ($p < 0.01$) and significant in the central region ($p < 0.05$). The reductions in error rates for Formant 2 were also more substantial than they were for Formant 1. This is reflected in the computed average errors, where it is seen that the percent-average errors and the percent-rms errors for Formant 2 in the initial region are greater than they are for Formant 1 before processing, and that they are lower after processing. **Figure 5** shows the effect of processing on the mean error for each formant, test condition, and error type, with the reductions in mean error rates expressed as percentages of the mean error in the unprocessed condition. It is apparent that the greatest improvements were for Formant 2 in the initial regions of the test words.

It is of interest to compare these results with the pattern of reduction of phoneme identification errors. **Table 4, Table 5, Table 6** and **Table 7**, which are based on data obtained in the experiments described in Hochberg, et al. (2), present average percent recognition scores before and after noise reduction processing for each of the 30 phonemes used in the AB word lists at a signal-to-noise ratio of 10 dB. Increases in scores for stop consonants are highly significant, with scores increasing from an average of 17.6 percent before processing, to an average of 29.2 percent after processing. By comparison, the changes in the scores for fricative consonants are small (2 percentage points, on average) and were not statistically significant. The scores for nasal and liquid consonants also did not show a statistically significant change (6 percentage points).

DISCUSSION

The Nucleus-22 cochlear implant processor estimates the frequencies of Formants 1 and 2. When noise is added to speech, these estimates deviate substantially from those that would have been obtained in the absence of noise. INTEL processing of noisy inputs reduces the magnitude of these deviations for Formant 2. Formant 2 is very important for the perception of speech by users of this type of cochlear implant. The first version of the Nucleus-22 system encoded only pitch frequency and Formant 2 amplitude and frequency data (F0F2A2) as compared with pitch, Formant 1, and Formant 2 frequency values (F0F1F2) encoded in the current version. However, analyses of the amount of information transferred by each of these coding schemes (8) show that the original scheme transmitted, on average, 50 percent of vowel information and 40 percent of consonant information, as contrasted with 63 percent and 48 percent, respectively, for the current coding method. Thus, it appears that most of the useful information for vowel and consonant identification is conveyed by Formant 2. Therefore, it is highly probable that the reduction in the deviations of Formant 2 frequency resulting from the INTEL processing of speech at a signal-to-noise ratio of 10 dB is the reason for the observed increase in subject test scores at that signal-to-noise ratio, as reported by Hochberg, et al. (2). The data presented there show that the average percent phoneme recognition for 10 cochlear implant users shifted upward by an amount corre-

Journal of Rehabilitation Research and Development Vol. 30 No. 1 1993



Figures 5a–5j.
Smoothed estimated frequencies of Formants 1 and 2 for 10 test words before and after INTEL processing.

Table 1.
Average percent absolute difference error in the measurement of Formant 1 and Formant 2 within the initial and central regions of each test word.

	Initial Region				Central Region			
	Formant 1		Formant 2		Formant 1		Formant 2	
Word	**Unpr**	**Proc**	**Unpr**	**Proc**	**Unpr**	**Proc**	**Unpr**	**Proc**
Bath	32	21	18	8	19	14	29	10
Dig	1	0	56	17	1	0	8	13
Five	22	2	48	4	13	15	17	19
Hum					44	26	46	23
Joke	2	1	62	10	5	3	88	45
Noose	182	109	53	13	33	47	87	57
Pot	40	45	18	18	31	41	21	13
Reach	1	1	52	15	18	11	10	11
Shell	21	45	9	9	17	18	37	31
Ways	1	1	44	16	11	7	9	13
Mean Error	34	25	40	12	19	18	35	24
LS	0.234		0.002		0.313		0.043	

Signal-to-noise ratio was 10 dB at the input to the INTEL noise suppression processor for both unprocessed (Unpr) and processed (Proc) conditions. Level of significance (LS) was computed by the Wilcoxon matched-pairs signed-ranks method.

Table 2.
Average percent rms difference error in the measurement of Formant 1 and Formant 2 within the initial and central regions of each test word.

	Initial Region				Central Region			
	Formant 1		Formant 2		Formant 1		Formant 2	
Word	**Unpr**	**Proc**	**Unpr**	**Proc**	**Unpr**	**Proc**	**Unpr**	**Proc**
Bath	42	25	20	9	25	20	35	12
Dig	1	0	77	21	2	1	11	15
Five	24	2	50	4	15	18	21	11
Hum					67	45	48	28
Joke	2	1	80	10	8	7	107	58
Noose	184	110	63	15	52	67	105	82
Pot	40	50	21	18	55	57	24	16
Reach	1	1	80	23	26	16	13	16
Shell	29	65	11	10	24	27	48	44
Ways	1	1	65	18	22	18	13	16
Mean Error	36	28	52	14	30	28	43	30
LS	0.234		0.002		0.278		0.042	

Signal-to-noise ratio was 10 dB at the input to the INTEL noise suppression processor for both unprocessed (Unpr) and processed (Proc) conditions. Level of significance (LS) was computed by the Wilcoxon matched-pairs signed-ranks method.

Journal of Rehabilitation Research and Development Vol. 30 No. 1 1993

Table 3.
Maximum absolute difference error in the measurement of Formant 1 and Formant 2 within the initial and central regions of each test word.

	Initial Region				Central Region			
	Formant 1		Formant 2		Formant 1		Formant 2	
Word	Unpr	Proc	Unpr	Proc	Unpr	Proc	Unpr	Proc
Bath	243	143	577	353	320	279	915	740
Dig	11	4	1051	795	32	18	505	719
Five	243	41	637	85	223	273	887	557
Hum					380	278	1206	719
Joke	16	9	910	146	196	115	1607	1033
Noose	412	233	1019	585	307	370	1592	1444
Pot	202	287	733	667	395	436	751	653
Reach	10	10	1496	563	250	189	969	189
Shell	172	385	747	410	333	350	1209	1301
Ways	8	7	1042	542	302	255	910	930
Mean Error	146	124	912	461	274	256	1055	829
LS	0.156		0.002		0.216		0.032	

Signal-to-noise ratio was 10 dB at the input to the INTEL noise suppression processor for both unprocessed (Unpr) and processed (Proc) conditions. Level of significance (LS) was computed by the Wilcoxon matched-pairs signed-ranks method.

Table 4.
Percent identification of vowels, diphthongs, and glides by 10 subjects using the Nucleus-22 cochlear implant.

	i	I	e	a	ae	^	u	eI	aI	^U	w
unprocessed	17	17	23	33	37	53	23	32	44	27	13
processed	42	42	35	32	45	42	52	35	61	48	32
difference	25	25	12	−1	8	−11	29	3	17	21	19
signed rank	9.5	9.5	5	−1	3	−4	11	2	6	8	7
tail probability		0.005									

Signal-to-noise ratio was 10 dB at the input to the INTEL noise reduction processor. Level of significance was computed by Wilcoxon matched-pairs signed-ranks method.

sponding to a 5-dB noise reduction when the noisy speech test materials were processed in the same manner as for this study (i.e., so as to improve the signal-to-noise ratio by an estimated 7 dB). While it is difficult to accurately measure the reduction of noise in the processed speech signals, the estimated reduction is consistent with the observed increase in subject performance.

This type of signal processing may also benefit other types of feature-extracting sensory aids in noisy conditions. The use of the INTEL method of noise reduction has been shown to improve the accuracy of spectrum-based speech recognition systems (9). Hence, it seems likely that the operation of cochlear implants that encode spectrum peaks or spectrum envelope characteristics also may be improved.

It is somewhat ironic to observe that cochlear implant users, who usually are severely to profoundly hearing-impaired, may in time be able to

Table 5.
Percent identification of nasals and liquids by 10 subjects using the Nucleus-22 cochlear implant.

	m	n	r	l
unprocessed	10	7	3	37
processed	16	29	10	26
difference	6	22	7	−11
signed rank	1	4	2	− 3
tail probability		0.31		

Signal-to-noise ratio was 10 dB at the input to the INTEL noise reduction processor. Level of significance was computed by Wilcoxon matched-pairs signed-ranks method.

Table 6.
Percent identification of fricatives by 10 subjects using the Nucleus-22 cochlear implant.

	f	s	sh	th	h	v	z
unprocessed	17	11	47	3	17	7	16
processed	13	3	61	3	29	10	13
difference	−4	−8	14	0	12	3	−3
signed rank	−3	−4	6		5	1.5	−1.5
tail probability		0.38					

Signal-to-noise ratio was 10 dB at the input to the INTEL noise reduction processor. Level of significance was computed by Wilcoxon matched-pairs signed-ranks method.

Table 7.
Percent identification of stop consonants by 10 subjects using the Nucleus-22 cochlear implant.

	p	t	k	ch	b	d	g	dj
unprocessed	7	13	20	43	17	27	7	7
processed	23	35	31	55	26	32	3	29
difference	16	22	11	12	9	5	−4	22
signed rank	6	7.5	4	5	3	2	−1	7.5
tail probability		0.008						

Signal-to-noise ratio was 10 dB at the input to the INTEL noise reduction processor. Level of significance was computed by Wilcoxon matched-pairs signed-ranks method.

function better in noisy environments than persons with mild to moderate hearing impairments. No single-channel method of attenuating wide-band random noise appears to be capable of consistently and/or substantially improving the intelligibility of noisy speech that is perceived through a conventional amplification system (10,11) whether by normal-hearing or by hearing-impaired listeners. On the other hand, the INTEL method is able to improve speech perception accuracy for cochlear implant users, and by an amount that corresponds to the apparent reduction in noise intensity. Because this

method is computationally intensive, there is little chance that it can be implemented in cochlear implant processors in the near future. However, similar benefits might be achieved by use of computationally less intensive transformation subtraction techniques, such as spectrum subtraction (12), which requires half the number of Fourier transform computations. If so, this would be a more practical approach to pursue.

ACKNOWLEDGMENTS

This work was supported in part by Grant PO1 DC00178-09 from the National Institute on Deafness and Other Communication Disorders.

REFERENCES

1. Skinner MW, Holden LK, Holden TA, Dowell R, Seligman PM, Brimacombe JA, Beiter AL. Performance of postlinguistically deaf adults with the wearable speech processor (WSP III) and mini speech processor (MSP) of the nucleus multi-electrode cochlear implant. Ear Hear 1991;12(1):3-22.
2. Hochberg IH, Boothroyd AS, Weiss M, Hellman S. Effects of noise suppression on speech perception by cochlear implant users. Ear Hear 1992;13(4):263-71.
3. Weiss M, Aschkenasy E. Automatic detection and enhancement of speech signals. Rome Air Development Center, Griffiss AF Base (NY): Report No: RADC-TR-75-77:1975.
4. Weiss MR, Aschkenasy E, Parsons TW. Study and development of the INTEL technique for improving speech intelligibility. Griffiss AF Base (NY): RADC-TR-75-108:1975.
5. Boothroyd A. Developments in speech audiometry. Sound 1968:2:3-10.
6. Widrow B, Glover JR, McCool JM, Kaunitz J, Williams CS, Zeidler JR, et al. Adaptive noise cancelling: principles and applications. Proc IEEE 1975;63(12):1692-1716.
7. Rabiner LR, Schaefer RW. Digital processing of speech signals. Englewood Cliffs (NJ): Prentice-Hall, 1978.
8. Blamey PJ, Dowell RC, Brown AM, Clark GM. Vowel and consonant recognition of cochlear implant patients using formant-estimating speech processors. J Acoust Soc Am 1987;82:48-57.
9. Cupples EJ, Foelker JL. Air Force speech enhancement program. Proceedings of Military Speech Tech-87. New York: Media Dimensions, 1987.
10. Lim JS, editor. Speech enhancement. Englewood Cliffs (NJ): Prentice-Hall, 1983.
11. National Research Council. Removal of noise from noise-degraded speech signals. Report of the Committee on Hearing, Bioacoustics and Biomechanics, Commission on Behavioral and Social Sciences and Education, National Research Council. Washington (DC): National Academy Press, 1989.
12. Boll SF. Suppression of acoustic noise in speech using spectral subtraction. IEEE Trans Acoustics, Speech, and Signal Processing 1979;27(2):113-20.