Effects of masking noise on vowel and sibilant contrasts in normal-hearing speakers and postlingually deafened cochlear implant users

Joseph S. Perkell[a)]
Speech Communication Group, Research Laboratory of Electronics, and Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Room 36-511, 50 Vassar Street, Cambridge, Massachusetts 02139 and Department of Cognitive and Neural Systems, Boston University, Boston, Massachusetts 02215

Margaret Denny
Speech Communication Group, Research Laboratory of Electronics, Massachusetts Institute of Technology, Room 36-511, 50 Vassar Street, Cambridge, Massachusetts 02139

Harlan Lane
Department of Psychology, Northeastern University, Boston, Massachusetts 02115 and Speech Communication Group, Research Laboratory of Electronics, Massachusetts Institute of Technology, Room 36-511, 50 Vassar Street, Cambridge, Massachusetts 02139

Frank Guenther
Department of Cognitive and Neural Systems, Boston University, Boston, Massachusetts 02215 and Speech Communication Group, Research Laboratory of Electronics, Massachusetts Institute of Technology, Room 36-511, 50 Vassar Street, Cambridge, Massachusetts 02139

Melanie L. Matthies
Department of Speech, Language and Hearing Sciences, Boston University, 635 Commonwealth Avenue, Boston, Massachusetts 02215 and Speech Communication Group, Research Laboratory of Electronics, Massachusetts Institute of Technology, Room 36-511, 50 Vassar Street, Cambridge, Massachusetts 02139

Mark Tiede
Speech Communication Group, Research Laboratory of Electronics, Massachusetts Institute of Technology, Room 36-511, 50 Vassar Street, Cambridge, Massachusetts 02139 and Haskins Laboratories, New Haven, Connecticut 06511

Jennell Vick
Department of Speech and Hearing Sciences, University of Washington, Seattle, Washington 98015

Majid Zandipour
Speech Communication Group, Research Laboratory of Electronics, Massachusetts Institute of Technology, Room 36-511, 50 Vassar Street, Cambridge, Massachusetts 02139 and Department of Cognitive and Neural Systems, Boston University, Boston, Massachusetts

Ellen Burton
Johns Hopkins School of Public Health and the Maryland Association of County Health Officers, Baltimore, MD 21205

(Received 28 February 2006; revised 9 October 2006; accepted 9 October 2006)

The role of auditory feedback in speech production was investigated by examining speakers' phonemic contrasts produced under increases in the noise to signal ratio (N/S). Seven cochlear implant users and seven normal-hearing controls pronounced utterances containing the vowels /i/, /u/, /ɛ/ and /æ/ and the sibilants /s/ and /ʃ/ while hearing their speech mixed with noise at seven equally spaced levels between their thresholds of detection and discomfort. Speakers' average vowel duration and SPL generally rose with increasing N/S. Average vowel contrast was initially flat or rising; at higher N/S levels, it fell. A contrast increase is interpreted as reflecting speakers' attempts to maintain clarity under degraded acoustic transmission conditions. As N/S increased, speakers could detect the extent of their phonemic contrasts less effectively, and the competing influence of

[a)]Electronic mail: perkell@speech.mit.edu

economy of effort led to contrast decrements. The sibilant contrast was more vulnerable to noise; it decreased over the entire range of increasing N/S for controls and was variable for implant users. The results are interpreted as reflecting the combined influences of a clarity constraint, economy of effort and the effect of masking on achieving auditory phonemic goals—with implant users less able to increase contrasts in noise than controls.
[DOI: 10.1121/1.2384848]

PACS number(s): 43.70.Mn, 43.70.Bk, 43.70.Dn, 43.66.Ts, 43.70.Fq [BHS] Pages: 505–518

I. INTRODUCTION

There is a substantial body of research concerning the effects of adverse speaking conditions on speech parameters. When the speaker perceives a deterioration in signal to noise ratio, either because of reduced signal levels or increased noise levels, that speaker will increase speaking sound level (Lane and Tranel, 1971; Van Summers *et al.*, 1988; Black, 1951; Hanley and Steer, 1949; Tartter *et al.*, 1993) and segmental duration (Van Summers *et al.*, 1988; Hanley and Steer, 1949). Utterances produced under such adverse conditions are more intelligible than those produced under optimal transmission conditions (Van Summers *et al.*, 1988; Dreher and O'Neill, 1958; Peters, 1955; Draegert, 1951).

These changes in sound level and durations under adverse conditions are consistent with those produced under instructions to speak clearly, as Lane *et al.* (1997) and Van Summers *et al.* (1988) noted. (There are, however, considerable differences among talkers: Hazan and Markham, 2004; Ferguson, 2004; Perkell *et al.*, 2002; Gagné and Tye-Murray, 1994.) Under clear speech instructions, vowel amplitudes and durations increase (Picheny *et al.*, 1986; Liu *et al.*, 2004). Furthermore, like speaking under adverse conditions, clear speech is also more intelligible than conversational speech (Picheny *et al.*, 1985, 1986; Chen *et al.*, 1983; Liu *et al.*, 2004; Krause and Braida, 2003; Payton *et al.*, 1994; Ferguson and Kewley-Port, 2002). Inference from similarities between clear speech and speaking under adverse conditions suggests that speakers may respond to clear speech instructions as though they were speaking under adverse conditions.

There is, however, an important difference between the changes in speech induced by instructions to speak clearly and those induced by adverse speaking conditions. Under clear speech instructions, phonemic contrasts are enhanced (Chen, 1980; Chen *et al.*, 1983; Moon and Lindblom, 1989; Picheny *et al.*, 1986). However, under more adverse speaking conditions, phonemic contrasts are characteristically degraded. For example, when the noise to signal ratio (N/S) is increased by subjecting normal-hearing speakers to loud masking noise, the speakers' vowel contrasts are reduced (cf. Bond *et al.*, 1989; Van Summers *et al.*, 1988). Likewise, if little or no signal can be heard, as in profound late-onset hearing loss, vowel contrasts (cf., Waldstein, 1990; Smyth *et al.*, 1991; Richardson *et al.*, 1993; Plant, 1984; Langereis *et al.*, 1997; Lane *et al.*, 2005) and sibilant contrasts (Lane and Webster, 1991; Matthies *et al.*, 1994) are also reduced compared to speakers with normal hearing (also see Kishon-Rabin *et al.*, 1999 on vowels).

The preceding considerations lead us to expect that speaking sound level will increase monotonically with N/S, whereas phoneme contrast distance will show an initial increase followed by a decline. These relations are schematized for vowels in Fig. 1. We define contrast distance for vowels as the average of Euclidean distances between all possible vowel pairs in acoustic (mel $1 \times$ mel 2) space. In Fig. 1(A), for speakers with normal hearing, the monotonic growth of SPL with N/S is shown by the solid line, while a hypothesized inflected phoneme contrast function is shown by a dotted-dashed line. We expect somewhat different contrast-distance functions in persons who have become profoundly deaf postlingually and then have had hearing partially restored with a cochlear implant. Their experience while deaf is likely to have led to reduced vowel contrasts (see above references). Moreover, the somewhat distorted hearing they receive from their cochlear implants may make them less able, compared to speakers with normal hearing, to use auditory feedback to help increase contrasts. Figure 1(B) schematizes hypothetical relations from cochlear implant users, with phoneme contrast at one-month postimplant shown by a dashed function, and at 1 year, by a dotted function. The differences between these two functions in overall levels of contrast and in the noise levels at which the inflection points occur reflect prior observations that experience with an implant can lead to contrast improvements (cf. Perkell *et al.*, 2001; Langereis *et al.*, 1997; Kishon-Rabin *et al.*, 1999). Note that with the present state of our knowledge we have no basis for specifying particular functional shapes for the three contrast-distance functions in Fig. 1—beyond the claims that the functions have (a) different overall levels, (b) a downturn, and (c) a certain ordering of that downturn for the three experimental conditions (controls, implant users at 1 month, and 1 year).

The kind of functional relations illustrated in Fig. 1 between N/S and phonemic contrast have not been reported previously, nor have speaking sound level and phonemic

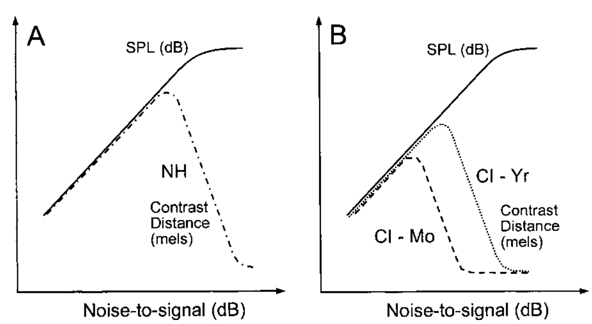

FIG. 1. Schematic illustration of hypotheses for normal-hearing speakers (panel A) and cochlear implant users (panel B) recorded at 1 month and 1 year post-implant.

contrast been examined together under this type of intervention.[1] In determining these functions for the cases of vowel and sibilant contrasts, the present experiment seeks to answer three general questions. First, we ask whether, in accord with Fig. 1, sound level functions are monotonic whereas contrast-distance functions are inflected, revealing a N/S threshold at which speakers typically stop increasing contrast and start reducing it. Second we ask what will be the behavior of speakers with late-onset deafness who have had some hearing restored recently by cochlear prosthesis. Several studies were cited above indicating that implant users produce reduced vowel contrasts. Consequently, the vowel contrasts produced by the implant users in this experiment may prove to be less robust in the face of increasing N/S than those produced by speakers with normal hearing. In other words, the function relating phoneme contrast to N/S may inflect at lower values of N/S for implant users than hearing controls. If so, will the function inflect at lower values in implant users when measured at 1 month than when measured after a year's experience with the implant (as postulated in Fig. 1)? Our third question is whether the findings for vowels and sibilants will differ from one another, and if so, how.

Clarity, effort, and the reduction in contrast: We speculate that the contrast distance functions schematized in Fig. 1 are the product of two underlying functions, one with positive slope, the other with negative slope, resulting in the inflection of the combined function. The function with positive slope would reflect a clarity effect—greater contrast distance under increasingly adverse conditions for communication. Underlying it is presumably the speaker's aim to communicate successfully despite increasing N/S. Underlying the negative slope—reduced phonemic contrast with increasing noise—is hypothetically a principle of least effort (Lindblom, 1990). As N/S increases, speakers are less and less able to hear the auditory consequences of their articulations, so without feedback their contrast distances fall, hypothetically due to a predominating effect of least effort. Consistent with these ideas, Perkell *et al.* (2002) found that when subjects were instructed to speak clearly some of them increased effort, as indexed by peak speed of articulation.

Economy of effort is a principle that guides speech motor control in the DIVA model of speech motor planning ("Directions into Velocities of Articulators;" Guenther, 1995; Guenther *et al.*, 1998, 2006; Perkell *et al.*, 2000), which we use as a framework for elaborating the current hypotheses and interpreting the results. In the model, phonemic goals are regions in auditory-temporal and somatosensory-temporal spaces. The goal regions are acquired and maintained with the use of auditory feedback, and speech movements from one goal to the next are programmed by feedback and feedforward subsystems. In the early stages of speech acquisition, feedback control predominates. As the sensory goals and feedforward commands are refined, movements become controlled mainly by the feedforward subsystem (although the feedback system remains available in case it is needed).

We noted above that when speakers encounter adverse speaking conditions (such as a noisy environment), they tend to speak more slowly, louder, and if possible, with increased contrast, in order to maintain intelligibility. Slower speech allows time for increased engagement of the feedback subsystem, which can help enhance or maintain contrast by providing information about the auditory consequences of the speech movements and allowing for feedback-based error correction if goals are not reached under feedforward control.

When hearing is lost in adulthood, the goal regions and feedforward commands may deteriorate gradually, leading to some diminution of phoneme contrasts. However, because the goals and feedforward commands are relatively robust and because somatosensory goals remain largely intact, basic phonemic identity is preserved.

The preceding reasoning leads to the prediction that under increasing N/S, vowel contrast will grow at the cost of expending more articulatory effort for as long as contrasts can be perceived by the speakers; then, as contrast perception and the use of auditory feedback control become more difficult, a threshold will be reached at which the balance begins to increasingly favor least effort and contrasts diminish (cf. Fig. 1).

Thus, the current study was designed to test the following hypotheses:

Hypothesis (1): In agreement with previous findings, over some range of increasing N/S values, speakers will not only speak more loudly and more slowly, they will also hyperarticulate, increasing contrast distance up to a point.

Hypothesis (2): At intermediate levels of noise, when speakers' auditory feedback of their own speech is masked to an extent that presumably begins to compromise their perception of phoneme contrasts, those speakers will continue to increase SPL toward a maximum level but their contrast distances will begin to fall.

Hypothesis (3): Contrast distance functions will inflect at higher N/S values for normal-hearing speakers than those for cochlear implant users, indicating that the normal-hearing speakers are less vulnerable to degradation of transmission conditions, presumably because they have robust and intact auditory goals and feedforward control mechanisms (cf. Guenther *et al.*, 1995, 2006; Perkell *et al.*, 2000) as well as better hearing.

Hypothesis (4): Implant users' contrast distance functions will inflect at higher N/S values after a year's implant use than after a month, indicating that their speech is less vulnerable to degradation of transmission conditions after experience with the implant—presumably because of retuning of auditory feedback and phonemic contrasts.

II. METHODS

A. Participants

The hypotheses set forth above were tested with two groups of participants. An experimental group comprised of five male and two female postlingually deaf, adult volunteers, who received cochlear implants at an average age of 52. The implant was either the Clarion (Advanced Bionics, CIS strategy; Wilson *et al.*, 1995) or the nucleus device (Cochlear Corporation; Blamey *et al.*, 1987; McKay and McDermott, 1993). The implant users were referred to our labora-

TABLE I. Participant characteristics.

1	Implant User ID	MM	MO	FJ	MK	FI	MJ	MP
2	Etiology	Noise (WWII)	Blood clot	Infection	Unknown	Auto-immune	Hereditary	Hereditary
3	Age at onset of hearing change	20	60	5	18	19	Early 30s	Birth
4	Age at onset of profound loss	72	67	45	28	54	43	26
5	Age at implantation	78	72	46	28	56	49	36
6	Hearing aid used pre-CI: L, R, both	Both	Left	None	None	None	Left	Both
7	Implant; ear	Clarion; R	Clarion C11 HiFocus; R	Clarion C11 HiFocus	Clarion Auria	Nucleus-24	Clarion C11 HiFocus	Clarion C11 HiFocus; R
8	Processing strategy: 1 month	CIS	CIS	SAS (8 chan.) Input dyn. range: 55 dB	HiRes-S	(not available)	CIS	CIS
9	Processing strategy: 1 year	CIS (no changes)	HiRes-P	HiRes-S Input dyn. range: 60 dB	HiRes-S (changed rate, pulse width)	SPEAK	HiRes-S	CIS (no changes)
10	Vowel perception (pre, year) % correct	32, 77	51[a]	26, 55	25, 95	29, 62	35, 96	27, 74
11	Consonant perception (pre, year) % correct	32, 50	49[a]	25, 57	24, 92	26, 44	21, 74	21, 74
12	Noise level: 1, 7; range, 1 mo. (dB)	61, 82; 21	45, 95; 50	38, 90; 52	39, 75; 36	51, 75; 24	61, 88; 27	40, 94; 54
13	Noise level: 1, 7; range, 1 yr. (dB)	53, 93; 40	53, 94; 41	70, 95; 25	51, 81; 30	59, 91; 32	59, 83; 24	73, 95; 22
14	Change in dynamic range (dB)	19	−9	−27	−6	8	−3	−32
15	Normal-Hearing Speaker ID	FNH5	FNH3	MNH2	FNH7	FNH6	FNH4	MNH1
16	Noise level 1, 7; range (dB)	23, 95; 72	17, 95; 78	23, 95; 72	23, 95; 72	23, 95; 72	15, 95; 80	21, 95; 74

[a]Data could not be obtained.

tory by the Massachusetts Eye and Ear Infirmary or the University of Massachusetts Memorial Medical Center and were paid for their participation. Table I presents pertinent characteristics of these individuals. Rows 10 and 11 of the table show that the implant users demonstrated substantial gains in vowel and consonant perception between measures made preimplant and at 1 year post implant.

A control group consisted of five female and two male paid volunteers, mean age 42, range 24–58, with no reported difficulties with speech or hearing. The control group initially consisted of four male and five female participants. Those over age 40 took a screening test, in order to determine approximate thresholds at 0.5, 1, 2, and 4 kHz. After a practice tone at 50 dB HL, sound pressure was increased in 5 dB increments from 0 to 25 dB HL. The series was presented twice at each of the four frequencies to each ear. Subjects who failed to report hearing the tone in any of the 16 series (2 ears, 4 frequencies, 2 trials) were excluded, which resulted in the exclusion of two males and left seven subjects in the control group. (On a test of phoneme recognition the remaining hearing participants scored at least 95% correct in vowel recognition and 90% correct in consonant recognition.) Consequently, the hearing controls are mainly female, while the implant users are mainly male. As explained below, in order to average data across subjects regardless of gender, and to compare patterns of averaged data between groups, the parameters of each speaker's vowel productions were converted to standard scores before averaging.

B. Procedures

1. Effects of masking noise

Two kinds of phoneme contrasts were examined, vowel and sibilant. The vowels /i/, /u/, /ɛ/ and /æ/ were elicited in the words *peet, poot, pet, and pat*, in the carrier phrase, "It's a _ please." The sibilants /s/ and /ʃ/ were elicited in the words *sot* and *shot*, in the carrier phrase "Say _ please." (The carrier phrase was changed for the sibilants to avoid subjects' tendency to elide the "a" in "It's a _.")

Implant users were recorded in two sessions following activation of the speech processors of their cochlear implants, at approximately 1 month postactivation and 1 year postactivation. For the normal-hearing controls, a single time sample was recorded. For all subjects, ten tokens of each of the four /pVt/ and the two /Sɑt/ ("S" = sibilant) words were elicited in random order within each of eight noise to signal levels. One N/S level had ambient noise only, which we call the "quiet" condition; the other seven had noise added at levels ranging between each subject's approximate thresholds for detection of the noise and for discomfort (described below). The presentations were blocked according to increasing levels of N/S.[2]

Participants were seated in a single-walled sound-attenuating booth (Eckel Industries) in a comfortable chair. A head-mounted electret microphone (Audio-Technica, model AT803B) was placed at a fixed distance of 20 cm from the participant's lips. The microphone was connected to a custom-built "feedback controller" (Technical Collaborative, Lexington MA), which mixed the subject's speech signal with specific calibrated levels of noise under computer control. The noise was approximately speech-shaped, with a spectral envelope that rolled off at 6 dB per octave.

Both subject groups received auditory feedback from the output of the feedback controller. For the implant users, the controller's output was connected to a laboratory speech processor, which was loaded with the settings currently in use on the subject's prosthesis. The implant users were allowed to adjust the overall gain to a level that felt "normal" to them before the recording commenced. For the normal-hearing participants, the device was connected to calibrated TDH-39 headphones with ear cushions and the volume set at a comfortable level for the subject.

The following procedure was used to establish an approximate "dynamic range" for each subject. In each session, for most of the subjects, the lowest noise level was determined by gradually increasing the noise level generated by the device until it was just detectable by the subject. The highest level was set to be just below that considered by the subject to be uncomfortable (see lowest and highest levels in Table I, rows 12, 13, and 16).[3] The same upper limit was used for all normal-hearing subjects, 95 dB SPL. For them as for the implant users, the five intermediate noise levels added were set at equal increments of SPL between the lowest and highest levels.

To provide a reference sound level for calculation of the sound pressure level of the subjects' produced vowels, a calibration signal was recorded while the subject remained silent. The signal was generated by an electrolarynx (Cooper-Rand Sound Source; Luminaud, Inc.; Mentor, OH) placed in front of the speaker's lips while an experimenter observed the sound pressure level on a sound level meter (C scale) placed next to the microphone.

For the recording, the subject's speech and the output of the feedback controller (containing the subject's speech mixed with noise, called the "mixed signal") were low-pass filtered at 7.2 kHz. The resulting signals were digitized directly to computer disk, each at a 16 kHz sampling rate.

2. Loudness-target control experiment

The changes in the speaker's contrast distance due to clarity and masking effects are potentially confounded with any changes in contrast distance due to simply speaking louder. In order to control for the effects on contrast distance of speaking louder (cf. Pickett, 1956), most of the participants served in a control experiment in which they read the same words they had read in the masking experiment, and contrast distance was measured without masking noise. Each speaker was asked to reproduce four of the seven speech sound levels he or she had produced in the masking experiment (called "target levels").[4] There were 10 repetitions of each of the utterances at each of the four target levels. To guide the speaker in this reproduction, the recording software generated a real-time visual display of the subject's sound level in the form of a moving bar graph, with a 4 dB wide target region. The centers of the target levels displayed were derived from the subject's productions in the masking-noise experiment at noise-added levels 1, 3, 5, and 7. All the subjects were readily able to follow the instruction to keep the bar approximately within the target region. Aside from these conditions, the recording and data extraction procedures were the same as in the masking experiment.

C. Data extraction

1. Vowels

Working with a display of the digitized speech signal of each utterance, an experimenter placed markers at the start and end of the vowel in each /pVt/ token. The parameters F1 and F2 were extracted algorithmically (while monitored and corrected, if necessary, by the experimenter) from an LPC spectrum around mid-vowel. A 40 ms analysis window for F0 and a 25 ms window for the formants were used. The LPC filter order was chosen to optimize formant delineation for each subject. (For further details of procedures for formant extraction, see Lane *et al.*, 2005.) Vowel duration was calculated from the labeled start and end times, and SPL was calculated from the RMS over the entire vowel duration as a log ratio with the RMS of the calibration signal. Values in mels for each formant, M1 and M2, were calculated from the formula

$$M = 2595 \times L_{10}(1 + (F/700)).$$

An overall measure of vowel contrast, average vowel spacing (AVS), was calculated as the mean Euclidean distance separating members of all possible pairs of vowels in the $M1 \times M2$ space for each repetition, averaged across repetitions (Lane *et al.*, 2001).

To provide a basis for calculating the N/S for each subject's utterances, the SPL of the noise at each of the seven noise levels presented to that subject was derived from the RMS of a portion of the mixed signal recorded when the subject was not speaking (i.e., containing only the noise). Then, the N/S of each token was calculated as the difference (in dB) between the SPL of the noise and the SPL of that token's recorded speech signal.

2. Sibilants

The start and end of the sibilant frication noise were determined by visual inspection of the waveform and the spectrogram, and the spectral mean was extracted algorithmically at the mid-point of each sibilant.[5] A measure of sibilant contrast distance was calculated as the difference between the spectral means of /s/ and /ʃ/ for each repetition, averaged across repetitions (Matthies *et al.*, 1994; Perkell *et al.*, 2004).

III. RESULTS

A. Vowels

1. Normal-hearing speakers

Figure 2 shows plots of average values of the parameters AVS (labeled A, in mels), duration (D, ms) and SPL (S, dB) as a function of average (N/S, dB) obtained at each of the seven noise levels. The error bars show one standard error about the mean. For plotting purposes, the value of N/S for the quiet condition was arbitrarily set to be less than the

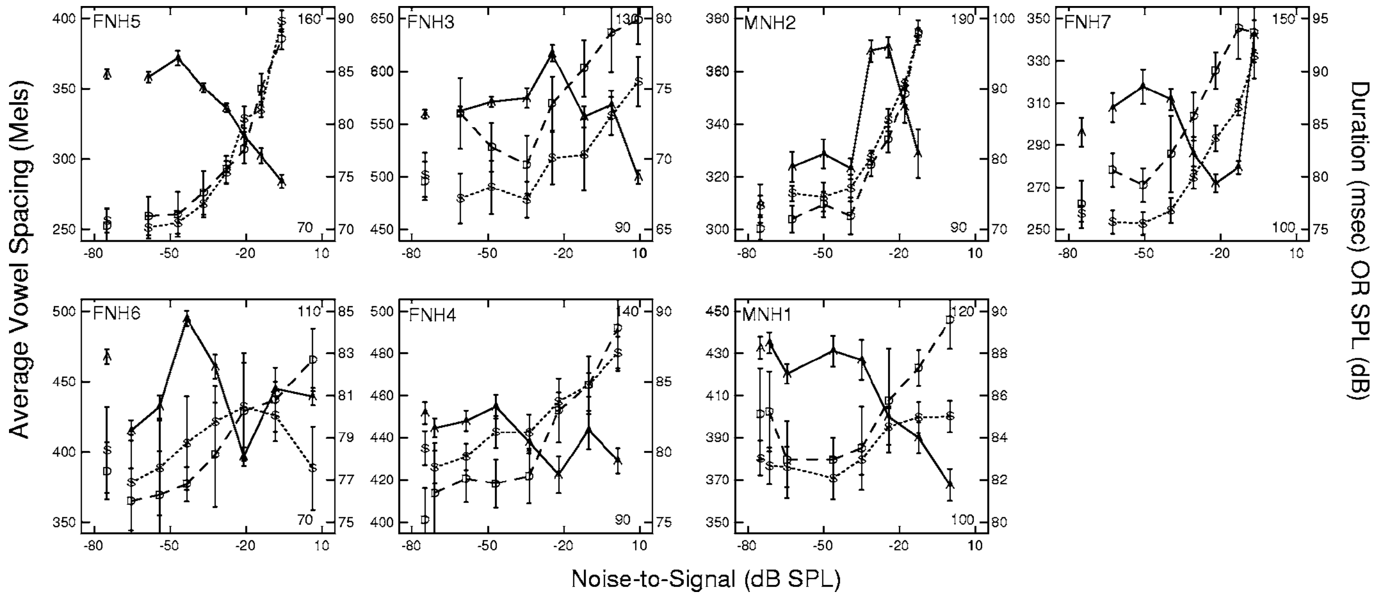

FIG. 2. AVS (A-mels), Duration (D-ms) and SPL (S-dB) vs. noise-to-signal ratio (dB) for each of the seven speakers who had normal hearing. For each panel, values of AVS are on the left vertical axis; the range of durations is shown by numbers in the upper and lower right-hand corners; values of SPL are on the right vertical axis. Values of N/S were obtained by multiplying extracted S/N by −1 and binning the resulting values into seven equally-spaced intervals. The values on the x-axis are located at the centers of the class intervals. Results from the quiet conditions are shown separately at the left of each plot (N/S is arbitrary.) Subject designations: H=hearing; M/F=male/female. NH=normal hearing. Error bars: standard error about the mean.

lowest value of N/S (from noise level 1). In each panel, values of AVS are displayed along the left vertical axis (mels); the range of durations is indicated by numbers in the upper and lower right-hand corners (ms), and values of SPL are displayed along the right vertical axis (dB). The first set of values in each panel is from the quiet condition and the remaining values, connected by lines, are from the seven N/S levels with added noise. Each panel shows data from a control participant with normal hearing. Note that the range of N/S is the same in all the panels, −80 to +10 dB. In order to show values for individual subjects and make the shapes of the functions as observable as possible, the scales of AVS, duration, and SPL vary across panels; they have been set so that the functions fill out the vertical space (which results in varying aspect ratios).

For speakers with normal hearing, the dynamic ranges displayed—the N/S range for each subject from his or her noise detection to noise discomfort levels—are roughly similar to one another, occupying most of the horizontal range in each panel. With a few exceptions and some variation among the subjects, the plots of each of the three dependent variables show similar trends across the subjects. SPL and duration grow approximately monotonically as N/S increases. The shapes of the AVS functions are more irregular; however most of them show an increase followed by a decrease. The most obvious exception to this pattern is AVS for subject FNH7 (top row, fourth panel), which increases substantially at the highest N/S level. Close examination of this subject's vowel spectra and listening to the utterances produced at the highest N/S level revealed that there was no discernable value of F2 for /u/ within its expected range (based on normative data from the literature) and the vowel sounded fronted, unlike an American English /u/. As a result, the algorithmically detected values of F2 were in the normal region of F3 for /u/, which caused inflated values of AVS. Subject FNH7 exhibited the same behavior in the loudness-target control experiment, so her data are not included in the plots or statistics of vowel data averaged across subjects (see below).

In order to examine group trends, Fig. 3 shows the data from six of the speakers with normal hearing (excluding FNH7) averaged across subjects. As noted above in Sec. II A, because subjects differed in the ranges of variables measured, values of each subject's dependent variables and N/S were rescaled by converting them to standard scores before averaging. Except for the quiet condition without

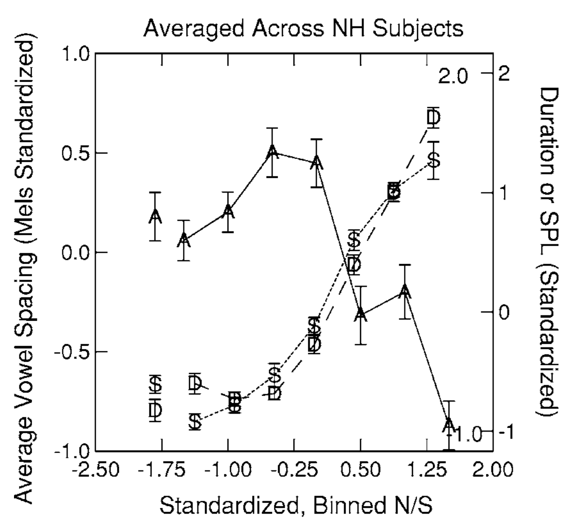

FIG. 3. Data averaged across six speakers with normal hearing (values in standard deviations). To compensate for inter-subject differences in data averages and ranges, each subject's data were standardized and the standardized values were averaged across subjects. See caption for Fig. 2 for remaining details. Data for Subject FNH7 were not included.

TABLE II. Summary of ANOVA results. $\eta^2 \times 100$=percentage of variance accounted for by the effect. η^2 (eta-squared) is calculated from $[F=(\eta^2/1-\eta^2)\times(\text{df error}/\text{df means})]$ (Young, 1993). NL: Noise level.

Row	Sample	Variable	Source	F	DF	p	$\eta^2 \times 100$
1	NH	AVSML	Subj	473	5, 33	<0.001	98
2			Noise level	19.5	6, 198	<0.001	37
3			NL×Subj	6.0	30, 198	<0.001	48
4			NL3>NL1	9.9	6, 33	<0.001	64
5			NL7<NL1	22.8	6, 33	<0.001	81
6		SPL	Subj	1130	5, 33	<0.001	99
7			NL	476	6, 198	<0.001	94
8			NL×Subj	92	30, 198	<0.001	93
9		DUR	Subj	128	5, 33	<0.001	95
10			NL	221	6, 198	<0.001	87
11			NL×Subj	23	30, 198	<0.001	78
12	CI—month	AVSML	Subj	156	6, 47	<0.001	95
13			Noise level	3.6	6, 282	<0.001	7
14			NL×Subj	1.6	36, 282	<0.01	17
15			NL7<NL1	3.6	7, 47	<0.001	34
16		SPL	Subj	251	6, 47	<0.001	96
17			NL	27.7	6, 282	<0.001	37
18			NL×Subj	4.9	36, 282	<0.001	38
19		DUR	Subj	328	6, 47	<0.001	97
20			NL	19	6, 282	<0.001	28
21			NL×Subj	13	36, 282	<0.001	62
22	CI—1 year	AVSML	Subj	410	6, 51	<0.001	98
23			Noise level	10	6, 306	<0.001	17
24			NL×Subj	3.5	36, 306	<0.001	30
25			NL4>NL1	3.8	7, 51	<0.05	34
26			NL7<NL1	16	7, 51	<0.001	69
27		SPL	Subj	1719	6, 51	<0.001	99
28			NL	176	6, 306	<0.001	78
29			NL×Subj	19	36, 306	<0.001	69
30		DUR	Subj	2146	6, 51	<0.001	99
31			NL	7.5	6, 306	<0.001	13
32			NL×Subj	8.0	36, 306	<0.001	48

added noise, the values of N/S were binned into seven equally spaced intervals, corresponding to the seven presented noise levels. The averaged standardized values of the dependent variables were plotted as a function of the average standardized value of noise level presented, expressed as the N/S class interval in which that noise level fell.[6] The centers of the seven N/S class intervals are shown on the abscissa in Fig. 3.

The resulting plot more clearly shows the trends observed in the individual data. SPL and duration grow monotonically as N/S increases. AVS increases then drops with increasing N/S. The local minimum at $N/S=0.50$ is due to downward fluctuations at the fifth point in the individual curves for FNH6, FNH3, and FNH4 (Fig. 2).

To test for the significance of observations made from this and subsequent plots of values averaged across subjects, two-way repeated-measures ANOVAs (with subjects as a category variable) and selected post hoc comparisons were computed on the original (unstandardized) data, separately for the normal-hearing speakers and for the implant users at 1 month and at 1 year. The ANOVAs were calculated on the unstandardized data to avoid possible violations of the assumptions underlying analysis of variance. The ANOVA results are presented in Table II and are indexed in the following text by the corresponding row numbers. The column labeled "$\eta^2 \times 100$" gives the percentage of variance accounted for by the effect. For the normal-hearing speakers and each of the three dependent variables, the effects of subject, noise level and their interaction were significant at $p < 0.001$ (AVS: Table II, rows 1–3, SPL: rows 6–8; Duration: rows 9–11). Values of $\eta^2 \times 100$ show that most of the effects accounted for large amounts of the variance, except for the effect of noise level and the interaction of subject with noise level on AVS (rows 2 and 3), possibly due to the irregular shapes of the individual AVS functions seen in Fig. 2.

Post hoc comparisons showed that the highest value of AVS (at the third noise level) was reliably greater than at the first noise level (Table II, row 4); AVS at the seventh noise level was reliably less than at the first noise level (row 5).

2. Cochlear implant users

Figure 4 shows individual plots for the seven implant participants. This figure is like Fig. 2, except in this case, there are two panels for each subject. The upper one is from data collected at 1 month postimplant; the lower, from data

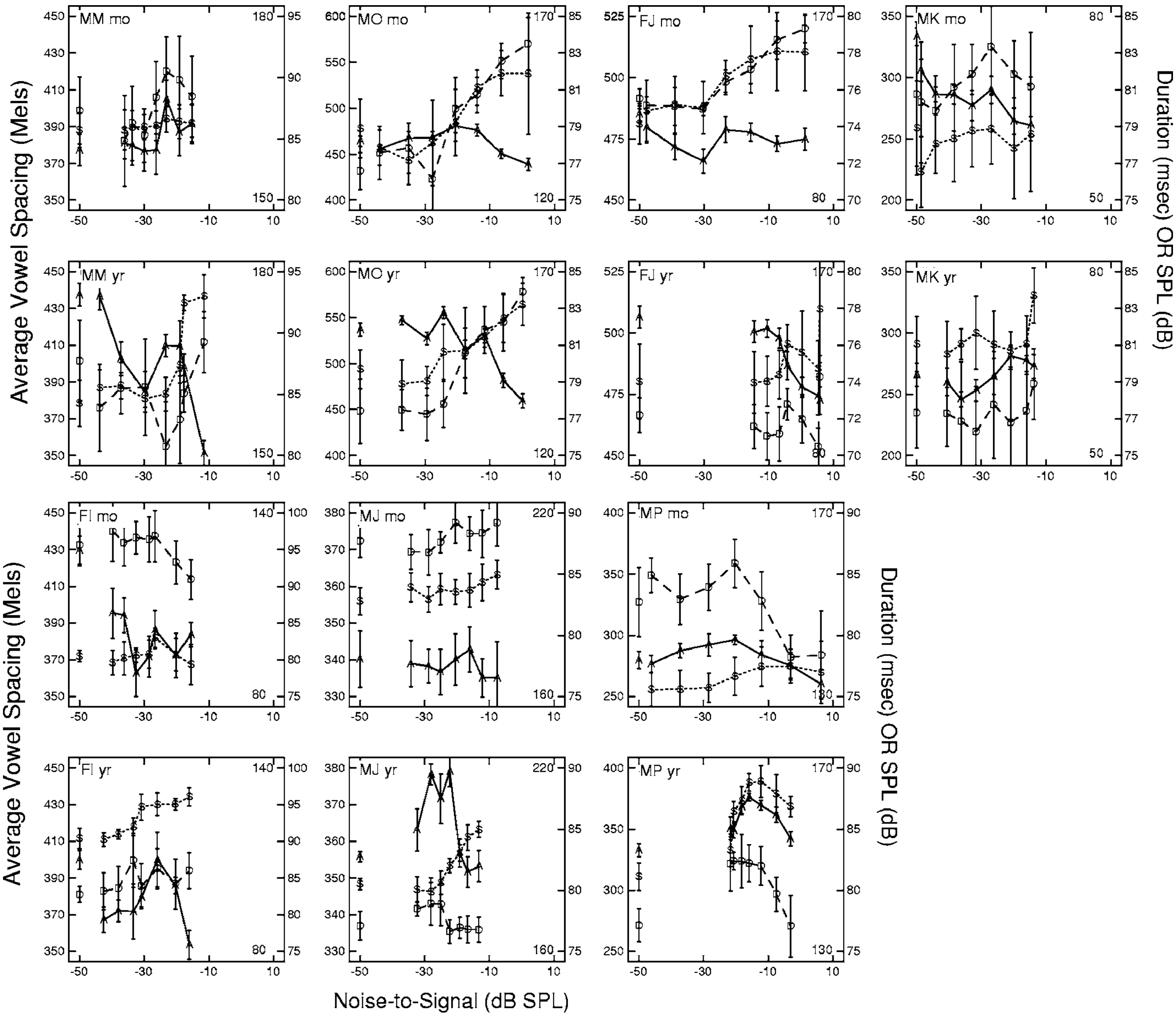

FIG. 4. AVS (A-mels), Duration (D-ms) and SPL (S-dB) vs. noise-to-signal ratio (N/S-dB) for each of the 7 implant users. There are two plots for each implant user, an upper one of data from one month and a lower one from one year post-implant. Subject designations: M/F=male/female. For further detail, see caption of Fig. 2.

collected at 1 year postimplant. The N/S range is the same in all the panels, −50 to 10 dB. The ranges of the dependent values are adjusted for each subject, but they are the same in each subject's 1 month and 1 year plots, to show changes between the two recording sessions.

The overall range of N/S displayed in the plots is 60 dB, 30 dB less than that for the speakers with normal hearing. This observation is consistent with the reduced dynamic range that implant users are known to have in general (cf. Hong *et al.*, 2003). For two of the implant users, FJ and MP, the dynamic range was considerably smaller at 1 year than 1 month (Table I, row 14) and the compression is in the direction of higher values. The opposite is true of subject MM, whose range is considerably greater at 1 year than 1 month (Table I, row 14). As can be seen from rows 8 and 9 in Table I, these changes cannot be accounted for by available information on implant processing strategies since MM and MP's strategies did not change between the two time samples.[7] Although there may be some overall regularity among the shapes of the functions for the dependent variables, the regularities are difficult to discern because of considerable difference among the subjects in the function shape (e.g., location of the inflection point) and sometimes large values of standard error about the mean.

Figure 5 shows data averaged across the seven implant users, at 1 month (panel A) and at 1 year (panel B) postimplant. As explained above, to combine individual speaker's data despite differences in the ranges of their variables, and to compare overall differences in levels of the dependent variables between 1 month and 1 year, before averaging over repetitions and speakers, all measures for each subject, including N/S, were converted to standard scores pooling across the two time samples. The functions in Fig. 5 show some similarities to those of the speakers with normal hearing in Fig. 3—more so at 1 year than at 1 month; however, there are also several differences.

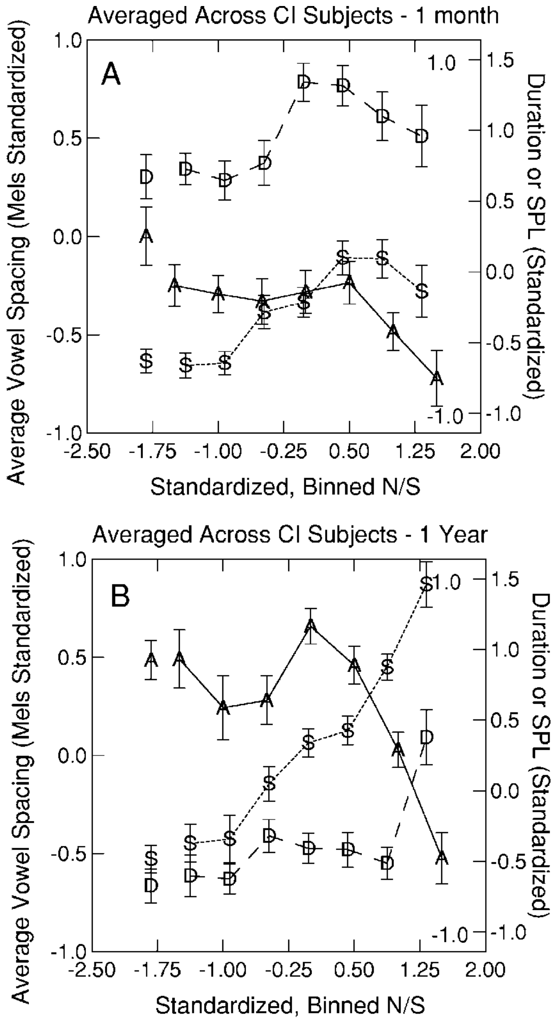

FIG. 5. Data averaged across the seven implant users, at one-month (panel A) and one-year (panel B) post-implant. See caption of Fig. 3 for remaining details.

a. AVS An ANOVA was computed on the unstandardized data for each time sample separately, with subject as a category variable and noise level as the treatment variable. The ANOVA results show main effects of subject and noise level on AVS at 1 month (Table II, rows 12 and 13) and at 1 year (rows 22 and 23) and a significant interaction between noise level and subject at both time samples (rows 18 and 24). In the 1 month data [Fig. 5(A)], AVS stays essentially flat with increased masking until it begins to drop at the sixth noise level. Effect sizes in Table II show that noise level accounts for 7% of the variance in AVS for the implant users at 1 month postimplant (row 13). At 1 year postimplant it accounts for more than twice as much of the variance (17%, row 23) but still less than for the controls (37%, row 2). AVS clearly begins at a higher mean standard score in the 1 year data [Fig. 5(B)] than in the 1 month data [Fig. 5(A)]. Post hoc comparisons showed that the highest AVS value in the 1 year data, at noise level 4, is slightly but reliably greater than its value at noise level 1 (row 25). AVS begins to fall at noise level 5 in the 1 year data [Fig. 5(B)] and reaches its lowest value at noise level 7, similar to the lowest value reached in the 1 month data (also at noise level 7). In both time samples, AVS is significantly lower at noise level 7 than at noise level 1 (rows 15 and 26).

b. SPL and duration As seen in Fig. 5, standardized durations are clearly shorter at 1 year than at 1 month postimplant, consistent with earlier observations on unstandardized measures (cf. Perkell *et al.* 1992). SPL grows with increasing N/S to a greater extent in the 1 year data than in the 1 month data. One-way repeated measures ANOVAs on the 1 month and 1 year data showed main effects of noise level on SPL and duration (one-month: rows 17, 20; one-year: rows 28, 31).

Turning to the subject variable and its interaction with noise level, there were significant effects of these two variables on SPL in both time samples (1 month: rows 16, 18; 1 year: rows 27, 29). SPL increased with noise level in both time samples, but the shape of the function resembles that of the normal-hearing speakers (Fig. 3) more in the 1 year data than in the 1 month data. There were also significant effects of subject and its interaction with noise level on duration in both time samples (1 month: rows 19, 21; 1 year: rows 30, 32).

B. Sibilants

Figure 6 shows plots of sibilant contrast distance (C), spectral mean for /s/ (S) and spectral mean for /ʃ/ (H) as a function of N/S (standardized) values. The N/S values are taken from the vowel data described above. The data are averaged across the seven control subjects [Fig. 6(A)] and the seven implant users at 1 month [6(B)] and 1 year [6(C)]. For the controls, contrast distance shows an initial nonsignificant rise from the quiet condition and then drops steadily as N/S increases. Observation of spectral means reveals that the decline in contrast distance is due mainly to an increase in the spectral mean for /ʃ/, beginning at the fourth noise level. The spectral mean for /s/ remains relatively flat across all N/S levels.

The data averaged across the implant users [Figs. 6(B) and 6(C)] are much more variable and show only one discernable trend: at 1 month postimplant [Fig. 6(B)], there is an initial increase in contrast distance followed by a decline over the first four N/S levels which appears to be due to a corresponding increase in spectral mean for /ʃ/. This contrast-distance decline in the 1 month data begins at a higher value than that found at any of the N/S levels in the 1 year data; at the end of the decline, it is in the same range as the values shown at 1 year across all N/S levels.

C. Control experiment

For each of the speakers who participated in the control experiment, whatever their hearing status, the sound levels they produced had no reliable effect on their vowel contrast or sibilant contrast. Between groups, implant users did not differ reliably from hearing speakers in vowel contrast ($F[1,88]=0.7, p>0.05$) nor in sibilant contrast ($F[1,83]=0.57, p>0.05$). There was no significant interaction between hearing status and the effects of produced sound level on vowel contrast ($F[3,264]=2.2, p>0.05$).

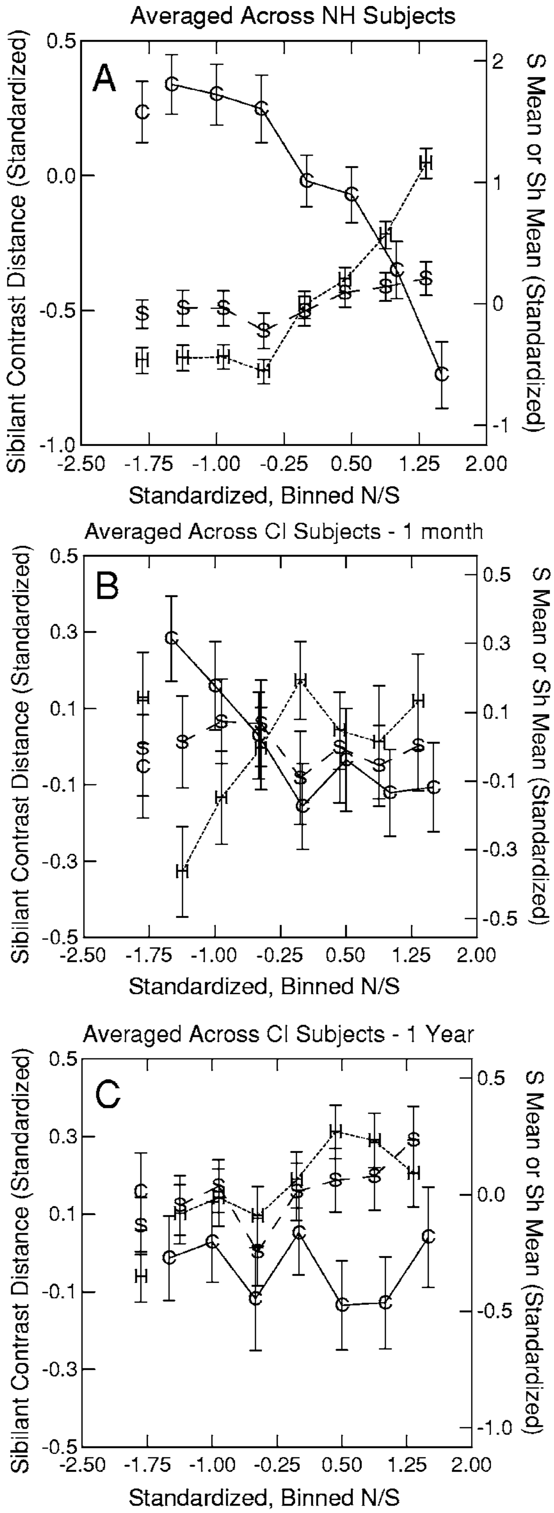

FIG. 6. Sibilant contrast distance (C), spectral mean for /s/ (S) and spectral mean for /ʃ/ (H) vs N/S (standardized values, as for Figs. 3 and 5). Results for speakers with normal hearing are shown in panel A, for implant users at one-month in panel B and for one-year post-implant in panel C. (For further details, see captions of Figs. 3 and 5.)

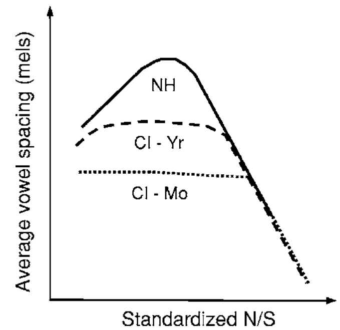

FIG. 7. Schematic diagram illustrating a possible explanation of the results. Average vowel spacing vs. normalized N/S. The solid function (NH) shows the result for the speakers with normal hearing; the dashed function (CI-Yr), for the implant users at one year post-implant, and the dotted line (CI-Mo), for the implant users at one month.

IV. DISCUSSION

A. Summary and interpretation of results

1. Vowels

Within each of the three experimental conditions (control, and implant users at 1 month, and 1 year post implant) there was considerable variation from one speaker to the next in the relations between the dependent variables and N/S. Among the implant users, possible sources of variance include processor strategy, insertion depth, and demographic factors identified in Table I, such as age at profound hearing loss. Furthermore, speakers with normal hearing differ in their understanding of how extensively they should adapt their speech when they are subjected to increased noise (Lane and Tranel, 1971). In spite of this between-speaker variation, when the results were averaged across subjects, they demonstrated support for the hypotheses. This support is most evident in the functions for AVS, SPL, and duration in the normal-hearing speakers (Fig. 3). Over the entire range of increasing N/S, vowel SPL and durations increase and at the lower N/S levels AVS begins to increase (Hypothesis 1). At intermediate noise levels, AVS begins to fall, while the other two parameters continue increasing (Hypothesis 2). At higher N/S levels, AVS drops to a level below the condition with no added noise. Thus we infer that as long as speakers can perceive the degree of vowel contrast in the presence of masking noise, they will attempt to increase it. The more masking interferes with the perception of vowel contrast, the more contrast drops. The results of the loudness-target control experiment—no systematic relation between AVS and produced sound level—indicate that the observed relations between vowel contrast and N/S are not simply due to speaking louder.

Hypothesis 3 predicted that the function would inflect at a higher N/S level for the speakers with normal hearing than for the implant users and Hypothesis 4 predicted that the inflection point would be at a higher N/S level for the implant users at one year than at one month. However, observation of Figs. 3 and 5 reveals a tendency toward the reverse ordering. The inflection point on the function labeled "A" is at the third or fourth noise level for the normal-hearing speakers, at the fourth level for the implant users at 1 year and at the fifth for the implant users at 1 month. A possible explanation for this preliminary observation, and for the observed differences in overall level of AVS among the three data sets, may be developed from further consideration of principles incorporated into the DIVA model as described in the introduction, along with the schematic diagram shown in Fig. 7.

The figure shows three hypothetical AVS functions: for normal-hearing speakers, (NH, solid line), for implant users at 1 year postimplant (CI-Yr, dashed line), and for implant users at 1 month postimplant (CI-Mo, dotted line). As schematized in the figure and observed above, the NH function rises and then drops. The CI 1 year function begins at a lower level of AVS than the NH function, and it only rises slightly before beginning to turn downward. The CI 1 month function begins at the lowest level and stays flat until it also turns downward. The observed tendency in ordering of the inflection points (third or fourth noise level for the normal-hearing speakers, fourth level for the implant users at 1 year, and fifth for the implant users at 1 month) is also approximated in the schematized functions in Fig. 7.

The introduction suggested that an inflected AVS function as exemplified by the solid (NH) curve in Fig. 7 could be the product of two underlying functions: (1) the speaker's increasing contrast distance as noise level increases and (2) a predominating influence of economy of effort as masking increasingly prevents the speaker from using auditory feedback to help achieve auditory goals. The contrasts produced by implant users at 1 month postimplant are presumably influenced by a long-term and gradual degradation of their feedforward commands and auditory goals while they were deaf. After implantation, a poorly tuned auditory feedback subsystem (which was still adjusting to the novel stimulus from the implant) could have made it impossible to increase vowel contrasts above current working levels. This would result in low levels of contrast despite increasing N/S as shown by the dotted (CI-Mo) function. After a year's experience with the implant, when the auditory feedback subsystem presumably had been tuned up and auditory goals and feedforward commands had been updated, the speakers operated at a higher level of contrast and were able to increase it somewhat under moderately adverse listening conditions, reflected in the dashed function (CI-Yr). In effect, a year's experience of refining auditory goals and retuning feedforward and feedback control subsystems could have raised the implant users' contrast "ceiling" to some extent, but not to the level of speakers with normal hearing.

Thus, all three data samples (normal-hearing speakers, implant users at 1 month and at 1 year) hypothetically may be characterized by a single function that incorporates the effects of clarity, economy of effort and masking, along with a superimposed clarity ceiling that presumably depends on the state of the speakers' feedback and feedforward control subsystems. These four factors taken together, then, would have established the observed ordering of inflection points, viz., at the third or fourth noise level for the normal-hearing speakers, at the fourth level for the implant users at one year, and at the fifth N/S level for the implant users at 1 month.

2. Sibilants

The speakers with normal hearing operated at a higher overall level of sibilant contrast (1134 Hz) than the implant users [669 Hz; ($F[1,198]=82$, $p<0.001$]. The functions in Fig. 6 show that sibilant contrast was more vulnerable to masking noise than the vowel contrast. Unlike AVS, sibilant contrast for the normal-hearing subjects showed no significant increase at lower values of N/S; it dropped continuously with increased levels of masking. Sibilant contrasts for the implant users were even more vulnerable to masking—showing large amounts of intersubject and within-subject variability and only a partially systematic relation to level of masking; a decline over the first four noise levels at 1 month postimplant. Considering that this decline is observed at 1 month, the lack of any systematic relation to masking at 1 year is somewhat puzzling. The sibilants' greater vulnerability to noise compared to vowels is not surprising. The sibilants are differentiated from one another by characteristics of their noise spectra, which are much more likely to be masked by speech-shaped noise than the higher amplitude spectral peaks of vowels that are being excited by a voicing source.

For the most part, the unstandardized contrast values in the quiet condition for the individual implant users (not shown) displayed generally good sibilant contrasts, although this relation in the average plots of standardized values relating sibilant contrast to N/S is inverted at 1 month [Fig. 6(B)] and is not significant at 1 month or at 1 year [Fig. 6(C)].

The finding that the decrease in sibilant contrast distance is due mainly to an increase in the spectral mean for /ʃ/ while that for /s/ remains flat [Figs. 6(A) and 6(B)] is compatible with earlier results. Matthies *et al.* (1996) describe a gradual increase in spectral median for /ʃ/ in an implant user during a period of about 17 min while his implant was turned off. The increase in spectral median was accompanied by a gradual movement of the tongue blade forward, which would diminish the size and raise the resonant frequency of the cavity anterior to the constriction. When the implant was turned on again, making it possible again for the subject to hear the consequences of his articulations, on the next utterance the subject had brought his tongue blade back to its normal place of articulation, returning the spectral median to its normal value.

Perkell *et al.* (2004) investigated the sibilant contrast further with measures of 19 normal-hearing speakers' produced spectral contrast distance, their use of contact of the tongue tip with the lower alveolar ridge (hypothetically a somatosensory goal), and their auditory acuity. The results showed that speakers who had higher auditory acuity for the sibilant contrast, and who used tongue contact for /s/ but not /ʃ/, produced the greatest contrast distances; those who evidenced moderate acuity or s-contact produced intermediate contrast distances; and those who showed neither high acuity nor s-contact produced the smallest contrast distances. These findings were interpreted as support for the hypothesis that articulatory goals are auditory and somatosensory: in this case, the goals would be a higher spectral center of gravity and tongue contact with the lower alveolar ridge for /s/ but not /ʃ/. The use of contact to stabilize the production of /s/ is one of a number of examples of saturation effects that stabilize the production of virtually any consonant (Perkell *et al.*, 2000, 2004). Thus the spectral mean for /s/ is relatively unaffected by increased masking, because speakers are able to rely on the saturation effect for that sound in order to maintain the articulation in a way that is not affected by the loss of auditory feedback. On the contrary, masking affects /ʃ/ presumably because speakers have to rely more on its auditory goal.

If speakers use an articulatory saturation effect for /s/ as described, the sibilant contrast would not be influenced by a tradeoff between clarity and economy of effort as much as the vowel contrasts would be. Were a speaker to change the amount of pressure between the tongue tip and lower alveolar ridge, that change in "effort" would presumably not be reflected in any change in the sibilant acoustics. Vowels, with their predominantly auditory/acoustic targets, generally are characterized by more continuous relations between changes in articulation and acoustics (except for place of articulation for /i/, /ɑ/ and /u/, Stevens, 1989, 1998; and degree of constriction for /i/, Perkell *et al.*, 2000; Perkell and Nelson, 1985).

B. Relations to other findings and conclusions

These findings are compatible with another experiment (Lane *et al.*, in press) conducted in our laboratory that employed the same implant users (plus an additional subject) and a group of normal-hearing subjects that included some of the same speakers as in this study. In the experiment reported by Lane *et al.*, AVS of eight vowels (as opposed to the current four) and sibilant contrast distance were examined preimplant and postimplant at two time samples—1 month and 1 year—with the implant processor (i.e., auditory feedback) turned off and on. The normal-hearing subjects, for feedback off, were presented with 95 dB of masking noise and for feedback on, no masking. In each of the recorded speech samples under feedback-on and feedback-off conditions, measures of both kinds of contrast, vowel and sibilant, were lower with feedback off than with it on. Parallel results were found in the current study: contrasts produced at the highest noise level (feedback maximally masked) were lower than the contrasts produced with no added noise. Also, contrasts for the larger vowel set were lower for the implant users at each time sample than were the contrasts of the normal-hearing speakers, as in the results of the present study.

In the only prior study that examined several vowel parameters produced under different levels of masking noise, Van Summers *et al.* (1988) measured RMS amplitudes, durations, F0, spectral slope and F1 and F2 from pronunciations of the digits "zero" through "nine" (plus some "control" words) by two male subjects with normal hearing. Five repetitions were pronounced at each of four levels of noise: quiet (no masking noise), 80, 90, and 100 dB. The two subjects differed somewhat from one another, but in general, amplitudes and durations followed the same trends as in the current study—increasing at higher noise levels. There was also a tendency for F0 values and spectral slope to increase with increasing levels of noise. F1 and F2 were analyzed separately and showed different trends for the two speakers; however, when we calculated AVS from the mean values in Figs. 6 and 7 of Van Summers *et al.* (1988, p. 922), both subjects showed lower values of AVS with 100 dB of masking noise than in the quiet.

Many investigators have found (e.g., Lane, 1963; Hawkins and Stevens, 1950; Hirsh *et al.*, 1954) that the intelligibility of speech mixed with noise decreases with increasing N/S. Assuming that speakers have implicit knowledge of noise-induced decrements in intelligibility, such knowledge provides motivation to enhance sound contrasts as much as possible in the face of environmental noise. The current results support the idea that speakers will increase clarity as much as they can, until the level of environmental noise begins to interfere with the use of auditory feedback mechanisms. With increasing noise, speakers are less and less able to perceive their own sound contrasts, and a presumed influence of economy of effort becomes a more predominant factor, causing contrast decrements. Speakers with compromised hearing (implant users in this case) habitually operate at lower levels of contrast than those with normal hearing, and when confronted with environmental noise, are less able to maintain and enhance contrasts.

ACKNOWLEDGMENTS

This research was supported by Grant No. R01-DC003007 from the National Institute on Deafness and Other Communication Disorders, National Institutes of Health. We are grateful to Dr. Donald Eddington of the Massachusetts Eye and Ear Infimary and Dr. Daniel Lee of the University of Massachusetts Medical Center for referring the implant users to the study, and to the implant users for their devotion of considerable amounts of their time. We also thank Advanced Bionics, Inc., and the Nucleus Corporation for their generous donations of research implant processors and three reviewers for their helpful comments on the manuscript.

[1]The study by Van Summers *et al.* (1988) is the closest in design to the current study. However, it was of only two male speakers with normal hearing; it did not employ as many noise levels and did not use levels near a detection threshold; and although F1 and F2 values were reported, contrast distances were not. Their results are compared further with the current study in the Discussion.

[2]In the earliest recordings, at 1 month postimplant for implant users MM, FJ, and MJ, the noise levels were randomized. The paradigm was changed to make the task more straightforward for the subjects.

[3]For all subjects, the noise was presented initially at 70 dB SPL, so that they would know what to listen for. To establish the lower threshold for all the implant users and all but three of the normal-hearing subjects, the noise was then presented at 11 dB SPL and the level was increased in 2 dB increments until the subject first reported hearing the noise. (For the first three normal-hearing controls, FNH3, FNH4, and MNH1, the lower threshold was set arbitrarily at 23 dB SPL.) When the lower threshold had been established for the implant users, the noise was presented again at 70 dB SPL and the level was increased in 2 dB increments until the subject indicated that the noise was as loud as he or she would tolerate. These procedures were carried out once on each subject. The upper limit was set at 95 dB SPL for all the normal-hearing controls. Although this method established only an approximation of each subject's dynamic range, it was considered adequate for the purposes of the experiment.

[4]The control experiment took place at a later date. Two of the seven NH subjects (FNH5, FNH6) and one of the seven implant users (MM) were not available to participate.

[5]Spectral mean has been shown to provide a robust, meaningful acoustic measure of the contrast (Forrest *et al.*, 1988; Jongman, Wayland, and Wong, 2000; Matthies *et al.*, 1994, 1996). Matthies *et al.* (1994) showed that late-deafened adults with reduced sibilant contrasts had improved contrasts after 6 months of implant use; spectral median and spectral skew gave similar results.

[6]The subject's produced vowel sound level varied somewhat from token to token within each of the presented noise levels. For those tokens produced with more extreme sound level values at any given noise level, the calculated value of N/S fell into the bin for the next higher or lower presented noise level; therefore, the number of tokens underlying the averaged values shown in Figs. 3, 5, and 6 vary somewhat from the standardized N/S value (bin) to the next.

[7]Differences in an implant subject's range between the thresholds for detection and discomfort at 1 month and 1 year may be due to the subjects' having used different sensitivity settings (for what felt "comfortable" at the time), which were beyond the experimental control.

Black, J. (**1951**). "The effect of noise-induced temporary deafness upon vocal intensity," Speech Monographs **18**, 74–77.

Blamey, P. J., Dowell, R. C., Brown, A. M., Clark, G. M., and Seligman, P. M. (**1987**). "Vowel and consonant recognition of cochlear implant patients using formant-estimating speech processors," J. Acoust. Soc. Am. **82**, 48–57.

Bond, Z. S., Moore, T. J., and Gable, B. (**1989**). "Acoustic-phonetic characteristics of speech produced in noise and while wearing an oxygen mask," J. Acoust. Soc. Am. **85**, 907–912.

Chen, F. R. (**1980**). "Acoustic characteristics and intelligibility of clear and conversational speech at the segmental level," Masters thesis, Massachusetts Institute of Technology (unpublished).

Chen, F. R., Zue, V. W., Picheny, M. A., Durlach, N. I., and Braida, L. D. (**1983**). "Speaking clearly: Acoustic characteristics and intelligibility of stop consonants," Working Papers, Research Laboratory of Electronics, Massachusetts Institute of Technology, **2**, 1–8.

Draegert, G. L. (**1951**). "Relationships between voice variables and speech intelligibility in high level noise," Speech Monographs **18**, 272–278.

Dreher, J. J., and O'Neill, J. J. (**1958**). "Effects of ambient noise on speaker intelligibility of words and phrases," Laryngoscope **68**, 539–548.

Ferguson, S. H. (**2004**). "Talker differences in clear and conversational speech: Vowel intelligibility for normal-hearing listeners," J. Acoust. Soc. Am. **116**, 2365–2373.

Ferguson, S. H., and Kewley-Port, D. (**2002**). "Vowel intelligibility in clear and conversational speech for normal-hearing and hearing-impaired listeners," J. Acoust. Soc. Am. **112**, 259–271.

Forrest, K., Weismer, G., Milenkovic, P., and Dougall, R. N. (**1988**). "Statistical analysis of word-initial voiceless obstruents: Preliminary data," J. Acoust. Soc. Am. **84**, 115–123.

Gagné, J. P., and Tye-Murray, N. (**1994**). *Research in Audiological Rehabilitation: Current Trends and Future Directions Monograph Supplement to Journal of the Academy of Rehabilitative Audiology* (American Academy of Rehabilitative Audiology, Cedar Falls), Vol. 27.

Guenther, F. H. (**1995**). "Speech sound acquisition, coarticulation, and rate effects in a neural network model of speech production," Psychol. Rev. **102**, 594–621.

Guenther, F. H., Ghosh, S. S., and Tourville, J. A. (**2006**). "Neural modeling and imaging of the cortical interactions underlying syllable production," Brain Lang **96**, 280–301.

Guenther, F. H., Hampson, M., and Johnson, D. (**1998**). "A theoretical investigation of reference frames for the planning of speech movements," Psychol. Rev. **105**, 611–633.

Hanley, T. D., and Steer, M. D. (**1949**). "Effect of level of distracting noise upon speaking rate, duration and intensity," J. Speech Hear Disord. **14**, 363–368.

Hazan, V., and Markham, D. (**2004**). "Acoustic-phonetic correlates of talker intelligibility for adults and children," J. Acoust. Soc. Am. **116**, 3108–3118.

Hawkins, J. E., Jr., and Stevens, S. S. (**1950**). "The masking of pure tones and speech by white noise," J. Acoust. Soc. Am. **22**, 6–13.

Hirsh, I. J., Reynolds, E. G., and Joseph, M. (**1954**). "Intelligibility of different speech materials," J. Acoust. Soc. Am. **26**, 530–538.

Hong, R. S., Rubinstein, J. T., Wehner, D., and Horn, D. (**2003**). "Dynamic range enhancement for cochlear implants," Otol. Neurotol. **4**, 590–595.

Jongman, A., Wayland, R., and Wong, S. (**2000**). "Acoustic characteristics of English fricatives," J. Acoust. Soc. Am. **108**, 1252–1263.

Kishon-Rabin, L., Taitelbaum, R., Tobin, Y., and Hildesheimer, M. (**1999**). "The effect of partially restored hearing on speech production of postlingually deafened adults with multichannel cochlear implants," J. Acoust. Soc. Am. **106**, 2843–2857.

Krause, J. C., and Braida, L. (**2003**). "Effects of listening environment on intelligibility of clear speech at normal speaking rates," Iranian Journal of Audiology **2**, 39–47.

Lane, H. (**1963**). "Foreign accent and speech distortion," J. Acoust. Soc. Am. **35**, 451–453.

Lane, H., Denny, M., Guenther, F. H., Matthies, M. L., Menard, L., Perkell, J. S., Stockmann, E., Tiede, M., Vick, J., and Zandipour, M. (**2005**). "Effects of bite blocks and hearing status on vowel production," J. Acoust. Soc. Am. **118**, 1636–1646.

Lane, H., Matthies, M. L., Denny, M., Guenther, F. H., Perkell, J. S., Burton, E., Tiede, M., Vick, J., and Zandipour, M. (submitted).

Lane, H., Matthies, M. L., Perkell, J. S., Vick, J., and Zandipour, M. (**2001**). "The effects of changes in hearing status in cochlear implant users on the acoustic vowel space and coarticulation," J. Speech Lang. Hear. Res. **44**, 552–563.

Lane, H., and Webster, J. (**1991**). "Speech deterioration in postlingually deafened adults," J. Acoust. Soc. Am. **89**, 859–866.

Lane, H., Wozniak, J., Matthies, M., Svirsky, M., Perkell, J., O'Connell, M., and Manzella, J. (**1997**). "Changes in sound pressure and fundamental frequency contours following changes in hearing status," J. Acoust. Soc. Am. **101**, 2244–2252.

Lane, H., and Tranel, B. (**1971**). "The Lombard sign and the role of hearing in speech," J. Speech Hear. Res. **14**, 677–709.

Langereis, M. C., Bosman, A. J., van Olphen, A. F., and Smoorenburg, G. F. (**1997**). "Changes in vowel quality in post-lingually deafened cochlear implant users," Audiology **36**, 279–297.

Lindblom, B. (**1990**). "Explaining phonetic variation: A sketch of the H & H theory," in *Speech Production and Speech Modeling*, edited by W. J. Hardcastle and A. Marchal (Kluwer, Dordrecht), pp. 403–439.

Liu, S., Del, R. E., Bradlow, A. R., and Zeng, F. G. (**2004**). "Clear speech perception in acoustic and electric hearing," J. Acoust. Soc. Am. **116**, 2374–2383.

Matthies, M. L., Svirsky, M. A., Lane, H. L., and Perkell, J. S. (**1994**). "A preliminary study of the effects of cochlear implants on the production of sibilants," J. Acoust. Soc. Am. **96**, 1367–1373.

Matthies, M. L., Svirsky, M. A., Perkell, J. S., and Lane, H. (**1996**). "Acoustic and articulatory measures of sibilant production with and without auditory feedback from a cochlear implant," J. Speech Hear. Res. **39**, pp. 936–946.

McKay, C. M., and McDermott, H. J. (**1993**). "Perceptual performance of subjects with cochlear implants using the Spectral Maxima Sound Processor (SMSP) and the Mini Speech Processor (MSP)," Ear Hear. **14**, 350–367.

Moon, S. J., and Lindblom, B. (**1989**). "Formant undershoot in clear and citation-form speech: A second progress report," Speech Transmission Laboratory, QPSR, 1/1989, 121–123.

Payton, K. L., Uchanski, R. M., and Braida, L. D. (**1994**). "Intelligibility of conversational and clear speech in noise and reverberation for listeners with normal and impaired hearing," J. Acoust. Soc. Am. **95**, 1581–1592.

Perkell, J. S., Matthies, M. L., Tiede, M., Lane, H., Zandipour, M., Marrone, N., Stockmann, E., and Guenther, F. H. (**2004**). "The distinctness of speakers' /s/-/ʃ/ contrast is related to their auditory discrimination and use of an articulatory saturation effect," J. Speech Lang. Hear. Res. **47**, 1259–1269.

Perkell, J. S., and Nelson, W. L. (**1985**). "Variability in production of the vowels /i/ and /a/," J. Acoust. Soc. Am. **77**, 1889–1895.

Perkell, J., Numa, W., Vick, J., Lane, H., Balkany, T., and Gould, J. (**2001**). "Language-specific, hearing-related changes in vowel spaces: A preliminary study of English- and Spanish-speaking cochlear implant users," Ear Hear. **22**, 461–470.

Perkell, J. S., Guenther, F. H., Lane, H., Matthies, M. L., Perrier, P., Vick, J., Wilhelms-Tricarico, R., and Zandipour, M. (**2000**). "A theory of speech motor control and supporting data from speakers with normal hearing and with profound hearing loss," Phonetica **28**, 233–272.

Perkell, J. S., Lane, H., Svirsky, M., and Webster, J. (**1992**). "Speech of cochlear implant patients: A longitudinal study of vowel production," J. Acoust. Soc. Am. **91**, 2961–2979.

Perkell, J. S., Zandipour, M., Matthies, M. L., and Lane, H. (**2002**). "Economy of effort in different speaking conditions. I. A preliminary study of inter-subject differences and modeling issues," J. Acoust. Soc. Am. **112**, 1627–1651.

Peters, R. W. (**1955**). "The effect of filtering of sidetone on speaker intelligibility," J. Speech Hear Disord. **20**, 371–375.

Picheny, M. A., Durlach, N. I., and Braida, L. D. (**1985**). "Speaking clearly for the hard of hearing I: Intelligibility differences between clear and conversational speech," J. Speech Hear. Res. **28**, 96–103.

Picheny, M. A., Durlach, N. I., and Braida, L. D. (**1986**). "Speaking clearly for the hard-of-hearing II: Acoustic characteristics of clear and conversational speech," J. Speech Hear. Res. **29**, 434–446.

Pickett, J. (**1956**). "Effects of vocal force on the intelligibility of speech sounds," J. Acoust. Soc. Am. **28**, 902–905.

Plant, G. (**1984**). "The effects of an acquired profound hearing loss on speech production," Br. J. Audiol. **18**, 39–48.

Smyth, V., Murdoch, B., McCormack, P., and Marshall, I. (**1991**). "Objective and subjective evaluation of subjects fitted with the cochlear multichannel prostheses: 3 studies," Australian Journal of Human Communication Disorders **19**, 31–52.

Stevens, K. N. (**1989**). "On the quantal nature of speech," J. Phonetics **17**, 3–46.

Stevens, K. N. (**1998**). *Acoustic Phonetics* (MIT Press, Cambridge, MA).

Tartter, V. C., Gomes, H., and Litwin, E. (**1993**). "Some acoustic effects of listening to noise on speech production," J. Acoust. Soc. Am. **94**, 2437–2440.

Van Summers, W., Pisoni, D. B., Bernacki, R. H., Pedlow, R. I., and Stokes, M. A. (**1988**). "Effect of noise on speech production: Acoustic and perceptual analyses," J. Acoust. Soc. Am. **84**, 917–928.

Waldstein, R. (**1990**). "Effects of postlingual deafness on speech production: Implications for the role of auditory feedback," J. Acoust. Soc. Am. **88**(5), 2099–2114.

Wilson, B., Lawson, D., Zerbi, M., Finley, C., and Wolford, R. (**1995**). "New processing strategies in cochlear implantation," Am. J. Otol. **16**, 669–681.

Young, M. A. (**1993**). "Supplementing tests of statistical significance: Variation accounted for," J. Speech Hear. Res. **36**, 644–656.