

April 01, 2007

## Time course of speech changes in response to unanticipated short-term changes in hearing state

Joseph S. Perkell

*Massachusetts Institute of Technology; Boston University*

Harlan Lane

*Northeastern University*

Margaret Denny

*Boston University; Massachusetts Institute of Technology*

Melanie L. Matthies

*Boston University; Massachusetts Institute of Technology*

Mark Tiede

*Massachusetts Institute of Technology; Haskins Laboratories, Connecticut*

*See next page for additional authors*

---

### Recommended Citation

Perkell, Joseph S.; Lane, Harlan; Denny, Margaret; Matthies, Melanie L.; Tiede, Mark; Zandipour, Majid; Vick, Jennell; and Burton, Ellen, "Time course of speech changes in response to unanticipated short-term changes in hearing state" (2007). *Psychology Faculty Publications*. Paper 21. <http://hdl.handle.net/2047/d20000868>

---

**Author(s)**

Joseph S. Perkell, Harlan Lane, Margaret Denny, Melanie L. Matthies, Mark Tiede, Majid Zandipour, Jennell Vick, and Ellen Burton

# Time course of speech changes in response to unanticipated short-term changes in hearing state

Joseph S. Perkell<sup>a)</sup>

*Speech Communication Group, Research Laboratory of Electronics, and Department of Brain and Cognitive Sciences, MIT, Room 36-511, 50 Vassar Street, Cambridge, MA 02139 and Department of Cognitive and Neural Systems, Boston University, Boston, Massachusetts 02215*

Harlan Lane

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

Margaret Denny

*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 Communication Disorders, Boston University, 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 06510*

Majid Zandipour

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

Jennell Vick

*Department of Speech and Hearing Sciences, University of Washington, Seattle, Washington 98015*

Ellen Burton

*Johns Hopkins School of Public Health and the Maryland Association of County Health Officers, Baltimore, Maryland 21205*

(Received 3 May 2006; revised 19 January 2007; accepted 19 January 2007)

The timing of changes in parameters of speech production was investigated in six cochlear implant users by switching their implant microphones off and on a number of times in a single experimental session. The subjects repeated four short, two-word utterances,  $/dV_1n\#SV_2d/$  ( $S=/s/$  or  $/ʃ/$ ), in quasi-random order. The changes between hearing and nonhearing states were introduced by a voice-activated switch at  $V_1$  onset. "Postural" measures were made of vowel sound pressure level (SPL), duration,  $F_0$ ; contrast measures were made of vowel separation (distance between pair members in the formant plane) and sibilant separation (difference in spectral means). Changes in parameter values were averaged over multiple utterances, lined up with respect to the switch. No matter whether prosthetic hearing was blocked or restored, contrast measures for vowels and sibilants did not change systematically. Some changes in duration, SPL and  $F_0$  were observed during the vowel within which hearing state was changed,  $V_1$ , as well as during  $V_2$  and subsequent utterance repetitions. Thus, sound segment contrasts appear to be controlled differently from the postural parameters of speaking rate and average SPL and  $F_0$ . These findings are interpreted in terms of the function of hypothesized feedback and feedforward mechanisms for speech motor control. © 2007 Acoustical Society of America. [DOI: 10.1121/1.2642349]

PACS number(s): 43.70.Mn, 43.70.Dn, 43.70.Bk, 43.66.Ts [BHS]

Pages: 2296–2311

## I. INTRODUCTION

Multiple parameters of speech production change when auditory feedback is lost or restored (Kishon-Rabin *et al.*, 1999; Lane and Webster, 1991; Waldstein, 1990; Cowie and Douglas-Cowie, 1983) or modified in some way (discussed below). Both the nature and the timing of these changes can shed light on the role of hearing in maintaining adult speech production. There is a considerable literature concerning the nature of changes in speech brought about by changes in auditory feedback, but less is known about the precise time course of such changes.

### A. Suprasegmental vs segmental changes

Speakers use auditory information to monitor listening conditions and adjust suprasegmental aspects of their speech accordingly. When listening conditions are degraded by the imposition of loud masking noise, speakers with normal hearing increase vocal amplitude (the Lombard effect-Lane and Tranel, 1971), fundamental frequency (Bond *et al.*, 1989; Clark *et al.*, 1987) and the duration of speech segments (Tartter *et al.*, 1993; van Summers *et al.*, 1988; Hanley and Steer, 1949). Similar changes have been observed in postlingually deaf cochlear implant users in whom auditory feedback has not been masked but blocked: temporarily by turning off their implant speech processors. This results in louder, slower speech (Svirsky *et al.*, 1992). Such changes tend to make speech more intelligible (van Summers *et al.*, 1988; Dreher and O'Neill, 1958; Peters, 1955; Draeger, 1951).

Another strategy that speakers might use to compensate for perceived degradations in speaking conditions would be to increase contrasts at the segmental level. Instead, decreases in vowel contrast have been observed when implant users' auditory feedback is blocked (cf. Perkell *et al.*, 2001) or the auditory feedback of normal-hearing speakers is masked (cf. Bond *et al.*, 1989; van Summers *et al.*, 1988). It appears contradictory that speakers would enhance some aspects of speech intelligibility while allowing others to degrade.

In search of the reason that speaking sound level and durations increase when auditory feedback is interrupted whereas vowel contrast decreases, Perkell *et al.*, (2007) exposed both speakers with normal hearing and those with cochlear implants to masking noise that ranged in intensity from just detectable to maximally tolerable. Measures included vowel duration, sound pressure level (SPL), and average vowel spacing (the mean separation of all possible vowel pairs in the formant plane; Lane *et al.*, 2001). Similar trends were observed for both subject groups, although the results from the implant users were more variable. Vowel duration and SPL increased with noise level as expected. Average vowel spacing also increased with noise intensity, but only for low to moderate noise levels. At the higher levels, vowel spacing declined. Perkell *et al.*'s interpretation

was that subjects tended to hyperarticulate their vowels for as long as they could hear the resulting contrast enhancement. As noise levels increased and they could no longer perceive vowel contrasts, an influence of economy of effort (Lindblom, 1990) prevailed, which resulted in the decreased vowel spacing at the highest noise levels.

This pattern of results suggests that listeners with normal hearing and postlingually deafened cochlear implant users respond similarly to degradation and to loss of auditory feedback and seek to increase their intelligibility with both suprasegmental and segmental cues, provided that they can hear themselves speak well enough. When they cannot, segmental contrasts decrease. This is true when (1) normal-hearing speakers speak in loud masking noise; (2) cochlear implant users speak in loud masking noise; and (3) cochlear implant users speak without benefit of prosthetic hearing. The difference between segmental contrasts on the one hand and durations and SPL on the other indicates that the two types of parameters might be controlled by somewhat separate mechanisms in which auditory feedback plays different roles.

The characteristics of these two types of parameters merits a brief discussion here. Based on a study of vowel production in cochlear implant users, Perkell *et al.* (1992) argued for the existence of a distinction between postural and segmental parameters. They suggested that speaking rate (which is reflected inversely in vowel durations), overall SPL, and average  $F_0$  are acoustic manifestations of postural settings that are adjusted rapidly to maintain intelligibility in the face of changing acoustic transmission conditions, for both normal-hearing speakers and users of cochlear implants. A possible alternative term for this class of speech parameters, "suprasegmental," usually refers to linguistically salient manifestations of prosody. Use of the term "postural" is intended to focus attention on nonlinguistic, relatively long-term average aspects of the speech signal. (The term postural was first introduced in this context by Stevens, Nickerson and Rollins, 1983). As to the segmental parameters, for each phonemic contrast in our research, a distance is calculated between the members of a contrast pair—which we call *contrast distance*. For example, the contrast distance of /a/-/ʌ/ is the mean separation (in Hz) in the  $F_1 \times F_2$  plane of tokens of those two phonemes; while the contrast distance of /s/-/ʃ/ is the difference in Hz between the average value of tokens of each phoneme in spectral median (Matthies *et al.*, 1994; 1996).

### B. Timing of postural changes

All of the effects on speech parameters discussed above have been observed in the course of single experimental sessions but those studies were not designed to determine how long it takes for a feedback-based compensatory response to emerge from normal variability in speech production. Svirsky *et al.* (1992) examined how quickly speech production parameters can change as a result of loss and restoration of auditory feedback. Three postlingually deafened cochlear implant users turned their speech processors off for 24 h prior to testing. On arriving in the laboratory, they read

<sup>a</sup>Author to whom correspondence should be addressed. Electronic mail: perkell@speech.mit.edu

words in carrier phrases, first with their processors off, then with their processors turned on. Their processors were then turned off again, and a final repetition of the reading task was performed. The investigators measured vowel duration, SPL, fundamental frequency ( $F_0$ ), and first and second formant frequencies.

In all cases, restoring hearing after 24 h of deprivation resulted in a significant decrease in SPL and  $F_0$ . The effects of turning the processor off again were inconsistent across subjects. Thus, compensatory adjustments in the postural parameters of sound level and  $F_0$  (cf. Perkell *et al.*, 1992) were more consistent when the speakers experienced a change from a nonhearing state to a hearing state than for a hearing change in the opposite direction. On the other hand, changes in durations (also a postural parameter) and segmental vowel formant frequencies were highly variable across subjects in both direction and time course. Svirsky *et al.* (1992) describe the time course of SPL changes as follows: “We can say with some certainty that SPL had reached a lower level by the first three to four occurrences of each vowel after turning the speech processor on, although inherent variability makes it difficult to indicate precisely when this change took place—it may have been well under way by the first or second occurrence of each token after turning the processor on” (p. 1290).

An important question raised by these observations is how long it takes for a compensatory adjustment to develop in response to a change in auditory feedback. In studies employing pitch-shifted feedback during the production of prolonged vowels, compensatory responses typically occur at a latency of 100–150 ms (Burnett *et al.*, 1997, 1998; Kawahara and Williams, 1996). Natke and Kalveram (2001) reported the effects of pitch shifts for randomly selected trials when normal-hearing subjects produced the nonsense word /tatatas/ with different stress patterns. The subjects’ voice  $F_0$  changed to compensate for the pitch shift in the first syllable, but only if that syllable was stressed. If the first syllable was unstressed, compensatory changes in  $F_0$  did not take place until the second syllable. The mean duration for stressed syllables was 325 ms and for unstressed syllables, it was 125 ms. Thus this result, indicating a delay of around 125 ms for feedback-based compensatory adjustments, is consistent with those reported by Burnett *et al.* (1997, 1998) and Kawahara and Williams (1996). Xu *et al.* (2004) also perturbed  $F_0$ , but in speakers of Mandarin, in which tone contours are used to differentiate CV words from one another. They found within-syllable compensatory responses with latencies as short as 100 ms. When contrasted with results showing longer delays, these findings were interpreted as indicating that the system for “regulation of voice  $F_0$  may be task dependent” (Xu *et al.*, 2004, p. 1168).

### C. Timing of segmental changes

Experiments in which speakers experienced unexpected changes in vowel formants in their auditory feedback throw further light on the timing of segmental vs postural parameters. Tourville *et al.* (2005) introduced abrupt unanticipated shifts of  $F_1$  (with an 18 ms delay) in the feedback of speak-

ers’ vowels embedded in /C $\epsilon$ C/ words. Subjects responded rapidly (within 100–200 ms) with partial compensatory changes in  $F_1$  (i.e.,  $F_1$  changes in the opposite direction). Such results demonstrate that the speaker is able under certain conditions to generate corrective motor commands during the current articulatory movement and that these changes can be observed experimentally if the movement lasts long enough. Purcell and Munhall (2006a) unexpectedly altered  $F_1$  in the auditory feedback of steady-state vowels and also found partial compensatory responses; however with a longer latency (up to 460 ms).

Most of the experiments cited above were designed mainly to look for closed-loop compensatory responses to modifications of auditory feedback—responses that are manifested during the production of the sound in which the modification is introduced. On the other hand, several recent “sensorimotor adaptation” experiments have been reported in which the focus was on compensation revealed in the production of subsequent sounds, and on “adaptation”—persistence of compensatory adjustments when auditory feedback is masked or when the perturbation is no longer present. In these studies, the experimental apparatus introduces incremental modifications of acoustic parameters in nearly real time and the subjects are unaware of the perturbations. Jones and Munhall (2002) introduced pitch shifts to speakers of Mandarin and found compensatory responses that persisted when the feedback was returned to normal. Houde and Jordan (1998, 2002) shifted  $F_1$  and  $F_2$  of vowels in whispered CVC (consonant-vowel-consonant) words to effectively change vowel quality that speakers heard themselves uttering. The speakers partially compensated for these changes in repeated elicitations, and the compensations persisted in the presence of masking noise. Using a similar technique, Villacorta *et al.* (2004, 2005) shifted vowel  $F_1$  in the feedback of speakers’ voiced VCV (vowel-consonant-vowel) utterances (also see Purcell and Munhall, 2006b). Partial compensatory adjustments were found; they persisted when feedback was masked and also for a short time after the perturbation was removed. The results of such studies may be interpreted as follows: speakers generate feedback-based error corrections, and if the movement lasts long, closed-loop corrections may be observed during the movement. Regardless of movement duration, the error corrections are incorporated into feedforward commands for the production of subsequent sounds (cf. Guenther *et al.*, 2006).

### D. Goal and hypotheses of this study

To our knowledge, there have been no studies designed to simultaneously examine segmental and postural responses to unexpected modifications of acoustic feedback in a way that could separate those responses from one another. The goal of the current study was to address this issue by examining how rapidly both kinds of parameters of speech production change when a speaker’s hearing is switched between blocked and unblocked states. The study uses repeated elicitations and multiple switches in hearing state, so that hearing-related effects can emerge from normal background variability in speech parameters. In order to examine the la-

TABLE I. Subject Characteristics. Male speakers are indicated by M, female, F. All but one of the subjects (FK) have participated in other studies in this laboratory. Subject identifiers are consistent across publications to facilitate comparisons across studies. (L designated left ear, R right.)

Speaker code	FI	FJ	FK	MM	MO	MP
Vowel perception	0.29	0.26	0.14	0.32	0.51	0.27
Consonant perception	0.26	0.25	0.31	0.32	0.49	0.21
Etiology	Auto-immune response	Infection	Infection	Noise (WWII)	Blood clot	Hereditary
Age at onset of change in hearing	19	5	48	20	60	Birth
Age at onset of profound loss	54	45	49	72	67	26
Age at cochlear implantation	56	46	50	78	72	36
Hearing aid used pre-CI: L, R, both	None	None	Both	Both	Left	Both
Implant: clarion/nucleus	Nucleus-24	Clarion	Nucleus-24	Clarion	Clarion	Clarion
Processor strategy	Spectral Peak Coding	Simultaneous Analog Stimulation	Advanced Combination Encoders	Continuous Interleaved Sampling	Continuous Interleaved Sampling	Continuous Interleaved Sampling

tency of compensatory responses with respect to the time of the change in hearing state, test utterances consist of sequences of two single-syllable CVC words and parameters in both words are examined. The study employs a group of postlingually deafened speakers who had had cochlear implants for one year. With implant users, it was possible to unexpectedly and completely block and restore prosthetic hearing by switching the input to their speech processors off and on at unpredictable intervals.

Based on the preceding background, the experiment is designed to test the following hypotheses:

1. Segmental contrasts will decrease when hearing is blocked, and they will increase when hearing is restored. Conversely, the postural variables (SPL,  $F_0$  and sound-segment duration) will increase when hearing is blocked, and they will decrease when hearing is restored.
2. If the switch in hearing state is introduced unexpectedly at the beginning of the vowel in the first of two CVC words spoken in sequence, changes in contrast distance and the postural variables will be evident in the second word, unless the duration of the vowel in the first word exceeds 150 ms. In that case, changes will be observed during that first vowel.
3. The latency of parameter changes following the switch will differ, depending on whether the parameter indexes a segmental or a postural variable.

## II. METHOD

### A. Subjects

Subjects were three male and three female postlingually deaf, adult, paid cochlear implant users. Informed consent procedures were carried out as approved by the Committee

on the Use of Humans as Experimental Subjects at MIT. The implant was either the Clarion (Advanced Bionics, Wilson *et al.*, 1995) or the Nucleus 24 device, (Cochlear Corp., Blamey *et al.*, 1987; McKay and McDermott, 1993). The subjects were referred by the Massachusetts Eye and Ear Infirmary or the University of Massachusetts Memorial Medical Center. Other studies from our laboratory (Lane *et al.*, in press; Lane *et al.*, 2005; Perkell *et al.*, 2007) include data from five of these individuals; to facilitate comparison of results, the same subject codes are used across all the studies. Pertinent characteristics of the subjects are summarized in Table I. To provide an approximate indication of subjects' ability to perceive speech, consonant and vowel identification scores are reported in Table I. All subjects participated in a vowel and consonant recognition test. Subjects were tested in a forced-choice task with eight vowels and another with 11 consonants. There were two recorded talkers, a male and a female, both speakers of Standard American English. With implant users' perception of their auditory feedback in mind, female subjects were presented recordings from the female talker and male subjects from the male talker. A practice test was given in which the stimuli were presented once each orthographically on a monitor and audibly with loudspeakers so the subject heard the stimuli knowing what the corpus would be. There were eight practice trials at the start of each session in which each of the syllables was presented once. During the practice, the listener was encouraged to adjust the volume control on the speakers to a comfortable loudness level. Then, subjects listened to four repetitions of each of three productions of each of the eight vowels for a total of 96 trials per block in random order; two blocks were presented in each testing session. (For more detail see Lane *et al.*, in press).

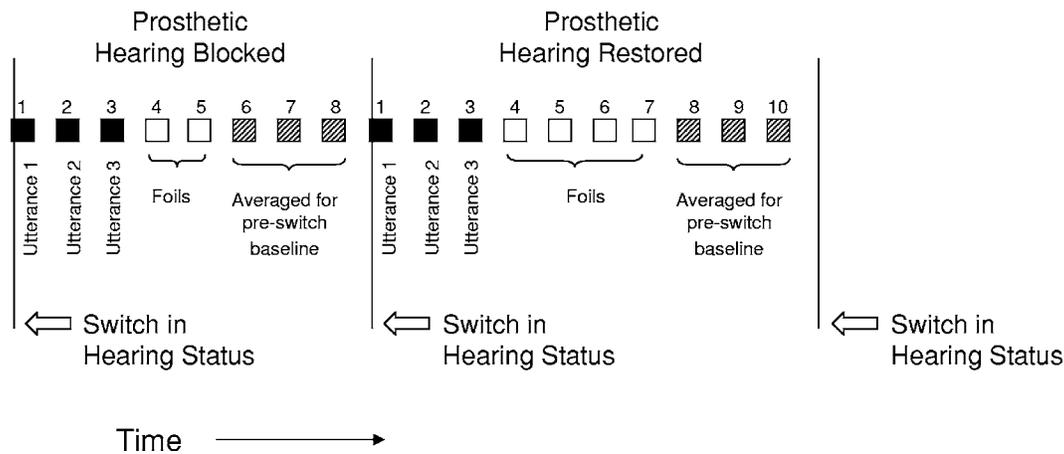


FIG. 1. Schematic of the experimental procedure. The squares represent individual two-word utterances (e.g., “Dun shed”), presented as a function of time. The vertical lines represent the switching of hearing state at 20 ms postvoicing onset in some utterances. Unfilled squares represent foils that were included to make the timing of the switches unpredictable, but were not analyzed; the number of foils varied from one to seven. Filled squares represent the utterances that were entered into the analysis. Results from the three utterances immediately prior to each switch (hatched squares) were averaged to obtain a pre-switch base line. For the three trials following a switch (solid black squares numbered from one to three) results were analyzed separately for each of the three postswitch utterances.

## B. Elicitation

The speech elicitation set consisted of four /dV1n#SV2d/ utterances of two words each: *Don shad*, *Don sad*, *Dun shed*, and *Dun said*. Thus, there were two vowel contrasts, /a/-/ʌ/ in the first word position and /æ/-/ɛ/ in the second, and the sibilant contrast /s/-/ʃ/ in the second. The subject was instructed to read the stimuli with neutral stress, at a comfortable loudness level. There was a brief practice period, in which the rate of stimulus presentation was adjusted to be comfortable for the subject, typically 3–3.5 s per stimulus item. The stimuli were presented in quasi-random order. In order to prevent the subject from anticipating the switches in hearing state, the number of stimulus presentations between switches was varied, with a minimum of eight items between switches.

Figure 1 illustrates the variable of utterance position; the term refers to the position in time of an utterance relative to a preceding or following switch in hearing status. Tokens with the same utterance position (and the same phonetic content) were averaged in order to determine, for example, the average duration of a vowel produced as the first utterance following a switch in hearing state from prosthetic hearing blocked to prosthetic hearing restored. Each square represents an utterance (e.g., *Dun shed*). Vertical lines represent switches in hearing status. The three solid black squares following each switch in hearing status are labeled utterance 1, utterance 2, and utterance 3. Blank squares represent foils, utterances that were not analyzed but were elicited to insure that subjects could not anticipate the timing of the switches in hearing status. The three utterances prior to a switch in hearing status (shaded squares) were averaged for a pre-switch base line.

A minimum of 15 repetitions of utterances containing each of the four vowels /a/, /ʌ/, /ɛ/ and /æ/ was recorded for each direction of switch (hearing blocked or restored) and for each of the three utterances immediately preceding and immediately following the switch. For example, the vowel /a/ in the first word in *Don sad* or *Don shad* was pronounced 15

times immediately following a switch that blocked hearing, 15 times as the second utterance following a switch that blocked hearing, and 15 times as the third utterance following a switch that blocked hearing. The need to provide multiple foils and to randomize vowels as well as sibilants resulted in a very large initial stimulus set. To avoid fatiguing the subjects, for the sibilants /s/ and /ʃ/, a minimum of 12 repetitions was recorded for each direction of switch and each of the three utterances immediately preceding and immediately following the switch in hearing state.

## C. Equipment

Each subject’s speech processor program currently in use was uploaded from his or her own processor, then downloaded to a laboratory-owned processor from the same manufacturer. Subjects then adjusted the processor controls until their own speech sounded “normal.”

The subject was seated in a sound-attenuating room in a comfortable office chair. A head-mounted electret microphone (Audio-Technica, model AT803B) was placed at a fixed distance of 20 cm from the subject’s lips and was connected through a preamplifier for recording the speech signal. A second microphone was placed near the subject’s ear. Its output was connected to a custom-built feedback controller (Technical Collaborative, Lexington, MA). The output of the controller was connected to the input of the speech processor, which in turn delivered the stimulation signal to the subject’s implant. The feedback controller included a voice-activated switch that turned the input to the speech processor on or off with a delay of 20 ms from the onset of a vowel. The feedback controller’s switching function was ramped to avoid the generation of abrupt changes in amplitude that would be heard as clicks. The stimuli were presented in text form on a computer monitor. Stimulus presentation and arming of the voice-activated switch were under computer control.

For calibration of sound pressure level, an electrolarynx (Cooper-Rand Sound Source; Luminaud, Inc.; Mentor, OH)

was 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. The calibration signal and the subjects' speech were low-pass filtered at 7.2 kHz and digitized in real time with a 16 kHz sampling rate.

#### D. Data extraction

Extracted data consisted of the postural variables of vowel  $F_0$ , SPL, duration, and the segmental variables of vowel  $F_1$  and  $F_2$  and sibilant spectral mean. Working with a display of the digitized speech signal of each utterance, an experimenter placed markers at the following points in each  $/dV_1n\#SV_2d/$  utterance: (1) at the onset of  $V_1$  ( $/a/$  or  $/A/$ ); (2) at the offset of  $V_1$ ; (3) at the onset of  $S$  (the sibilant  $/s/$  or  $/ʃ/$ ); (4) at the offset of  $S$ ; (5) at the onset of  $V_2$  ( $/æ/$  or  $/ɛ/$ ); and (6) at the offset of  $V_2$ . Vowel data were extracted from the 15 repetitions at each of the three utterances immediately preceding each switch in hearing state, and of each of the three utterances immediately following.

For the vowels,  $F_1$ ,  $F_2$ , and  $F_3$  were extracted algorithmically from an LPC (linear-predictive coding) spectrum around midvowel using a 25 ms analysis window. The LPC filter order was chosen to optimize formant extraction for each speaker. The algorithm displayed, for each vowel token, the initial measurements at the exact midpoint of the vowel; a broadband spectrogram on which was superimposed the formant trajectories that were detected; and, finally, the spectral cross section at the measurement time. If the first three formants were detected unambiguously in the regions expected for that vowel target, the experimenter accepted that token with those values. If not, the experimenter adjusted the measurement offset time slightly (as much as three glottal cycles in either direction) until the formants could be detected unambiguously.  $F_0$  was also estimated at the same offset over a centered 40 ms window, based on the filtered error signal autocorrelation sequence to minimize formant interaction (modified autocorrelation analysis; cf. Markel and Gray, 1976). Duration and rms amplitude were extracted algorithmically at the same time. The previously recorded calibration signal was used to convert the rms amplitude to dB SPL. Contrast distances between the vowels in the two first words ( $/a/$ ,  $/A/$ ) and between the vowels in the two second words ( $/æ/$ ,  $/ɛ/$ ) were calculated as Euclidean distances in the formant  $1 \times$  formant 2 plane expressed in mels.

Sibilant data were extracted from the 12 repetitions at each of the three utterances immediately preceding each switch, and from each of the three utterances immediately following. For each sibilant, the spectral mean was extracted algorithmically. Contrast distance between the sibilants was calculated as the average separation of the spectral means of the tokens of  $/s/$  and  $/ʃ/$  (Jongman, Wayland, and Wong, 2000; Forrest *et al.*, 1988; Matthies *et al.* 1994, 1996).

#### E. Graphing and statistical analyses

Figure 2 illustrates how the data were averaged over both utterance content and utterance position. As in Fig. 1, squares represent individual utterances and vertical lines rep-

resent switches in hearing status. The postural parameters of the second utterance following restoration of prosthetic hearing (number 2—black squares) have been selected for illustration. Each parameter mean is derived from data collected over the entire experimental session for a given utterance position and word, while the surrounding content varies unpredictably.

Repeated-measures analyses of covariance (ANOVAs) for each of the dependent variables were performed with subjects as the categorical variable. The effects of blocking and of restoring hearing were tested in separate ANOVAs. In order to test the hypotheses concerning parameter changes following switches in hearing status, six planned contrasts were evaluated for each dependent variable and each change in hearing state—blocked and restored. For each of the two vowel contrasts  $/a/$ - $/A/$  and  $/æ/$ - $/ɛ/$ , the mean contrast distance of the first, second and third utterances following the switch in hearing state were contrasted with the mean base line value, viz., the value averaged over the three utterances immediately preceding the switch. For the single sibilant contrast ( $/s/$ - $/ʃ/$ ), the effects of the switch on mean contrast distance in each of two vowel contexts ( $/æ/$  and  $/ɛ/$ ) were evaluated. For the three postural variables—duration, SPL and  $F_0$ —planned contrasts assessed the change from the three base line utterances pooled to the first, second, and third postswitch utterances; this was done for each of the two vowel positions in an utterance (first word or second word). Data for the two different vowels in each of those two positions were pooled (first word:  $/a/$ - $/A/$  and second word:  $/æ/$ - $/ɛ/$ ). The  $p < 0.01$  level of significance was adopted in testing these contrasts to reduce the chances of Type I errors.

### III. RESULTS

#### A. Contrast distance

##### 1. Vowels

Figure 3 illustrates the results of the analysis of vowel contrast distances, averaged across subjects. The vertical dashed lines represent the time of the switches, which actually occurred at the beginning of the vowel in the first of the two words in an utterance. The symbols to the left of these lines represent the parameter values averaged over the last three utterances prior to the switch. Vertical pairs of dotted lines represent the variable intervals in which stimulus items were presented, but not analyzed (called “foils” in the figure legend). For each panel, the question of interest is whether the results for each of the utterances labeled 1, 2, and 3 on the abscissa are different from the values measured in the pre-switch base line. Table II reports mean values and the outcome of statistical tests. Each row corresponds to a change in prosthetic hearing, an utterance position immediately following that change, and the vowel contrast embodied in the utterance. For each row, the table reports the average value of contrast distance in the three utterances preceding the switch, the value for the postswitch utterance, the  $F$  value of a planned comparison, and a measure of variance accounted for (eta-square). The changes in contrast distance following changes in hearing state were inconsistent and mostly quite small, median 1.6%. Six of the 12 contrasts

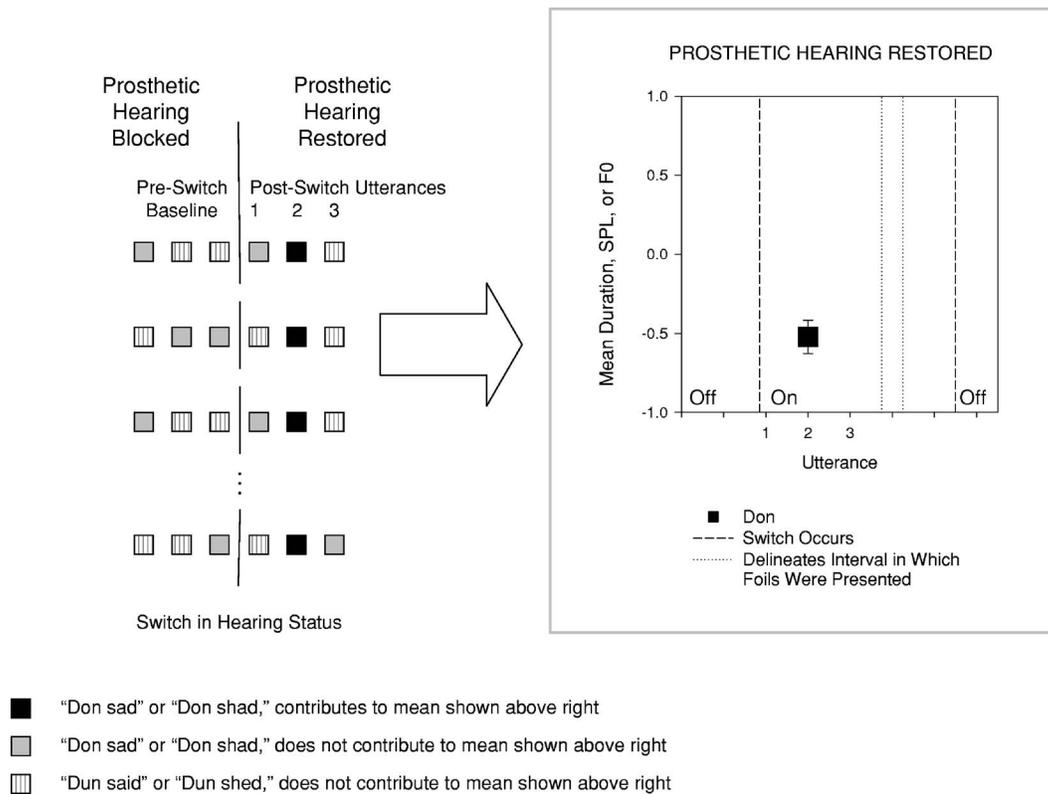


FIG. 2. Schematic illustrating how graphs of results were generated. The example shown here is for the /a/ in *Don* occurring as the second utterance after a switch in hearing status from prosthetic hearing blocked to restored. As in Fig. 1, each square represents an individual two-word utterance (e.g., *Don sad*). Solid filled boxes represent utterances of *Don* (followed by either *sad* or *shad*). Only the black (not the gray) utterances contribute to the data point shown in the graph at the right side of the figure. Hatched squares represent utterances of *Dun* followed by *said* or *shed*. In order to generate the single data point shown in the graph at the right, values were averaged across all 15 utterances of *Don* occurring during the second utterance after the switch in hearing status from hearing blocked to hearing restored. The vertical ellipsis indicates that not all 15 switches in hearing state are pictured here; the utterances shown would have been scattered in time throughout the experimental session.

were not statistically reliable at  $p < 0.01$ . Most changes were increases both when hearing was blocked and when it was restored (cf. Fig. 3). In an exception worth noting, after hearing was blocked, the /æ/-/ɛ/ vowel contrast declined 8 mels in the first utterance after the switch (from 90 to 82 mels, row 4), and when hearing was restored that contrast increased 5 mels (row 10). That change in contrast distance after blocking was not sustained in the second utterance (row 5) but did appear in the third. The 5 mel increase after restoring feedback was sustained in the second utterance (row 11) but not the third (row 12).

## 2. Sibilants

Sibilant results from subjects FI and FJ were excluded from analysis because some data were lost due to technical difficulties. The results from the remaining four subjects are illustrated in Fig. 4.

Mean sibilant contrast distances and the outcome of statistical tests appear in Table III. Again, dashed lines represent the occurrence of a switch and dotted lines the interval in which utterances were elicited but not analyzed (foils). Upper panels report results from hearing blocked, lower from hearing restored. Left-hand panels are for the vowel context /æ/, right-hand panels for /ɛ/.

Table III shows that changes in hearing state had no statistically reliable effects on sibilant contrast distance. This

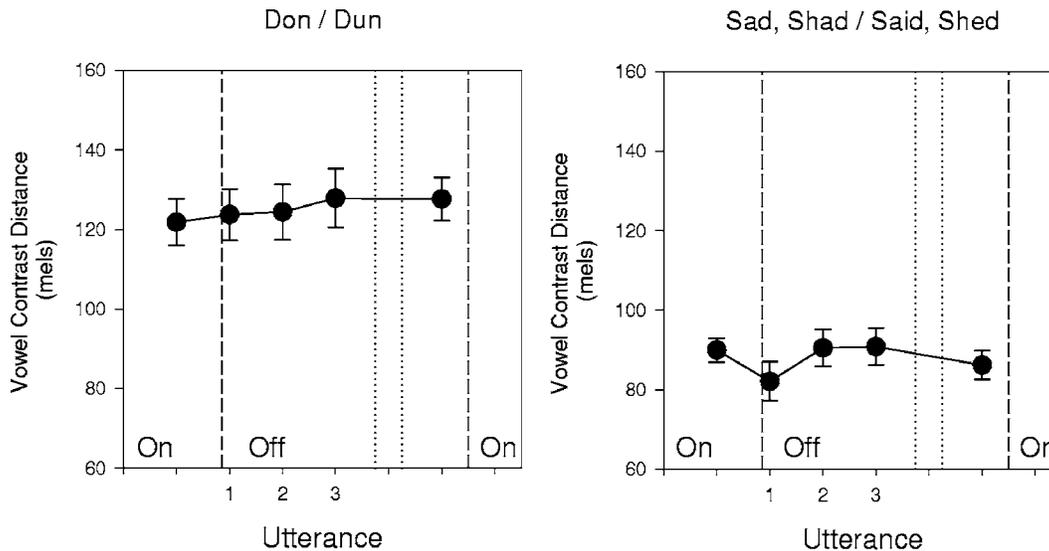
was true no matter whether postswitch utterances were considered singly or contrasted as a group with base line utterances. There is a drop worth noting in contrast distance on the first utterance in the /æ/ environment when prosthetic hearing was blocked (row 1 and upper left-hand panel). However, because of the large amount of variability, the drop is not reliable on this or the following two utterances (rows 2 and 3) which are much closer to the preswitch base line. This drop in contrast distance is not replicated in the /ɛ/ environment (there are increases on the second and third utterances, rows 5 and 6), nor is there evidence of an effect in either vowel context when prosthetic hearing was restored.

## B. Postural parameters

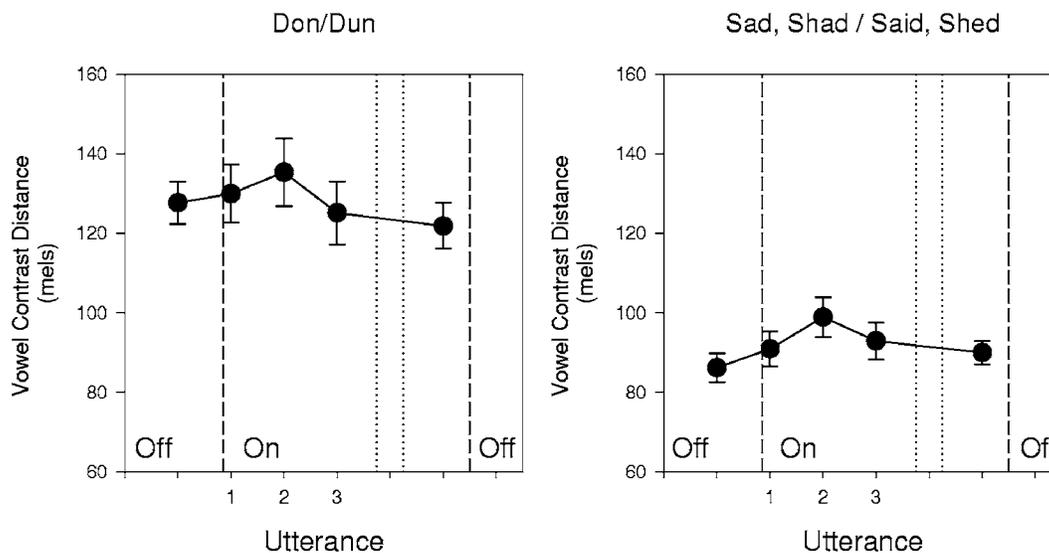
### 1. Duration

Figure 5 and Table IV summarize the effects of short-term changes in hearing state on vowel duration. Changes in hearing state were associated with large reliable changes in vowel duration. The changes in vowel duration following blocking or restoring of prosthetic hearing were all statistically significant at  $p < 0.01$  in each of the three postswitch utterances and both vowel positions (first word and second word of each utterance). When hearing was blocked, the rise in vowel duration from the base utterances to the first utterance postswitch was 12% for the mean of the two vowels

PROSTHETIC HEARING BLOCKED



PROSTHETIC HEARING RESTORED



----- Switch Occurs  
 ..... Delineates Interval in Which Foils Were Presented

FIG. 3. The effects on vowel contrast distance of blocking (top) and restoring (bottom) prosthetic hearing, as a function of utterance relative to a switch in hearing state, for vowels in the first word (left) and vowels in the second word (right). The vertical dashed lines represent the relative timing of the switches (which actually occurred at the beginning of the vowel in the first of the two words in an utterance; the line has been shifted for clarity). The symbols immediately to the left of these lines represent the average over three utterances immediately prior to the switch. Vertical pairs of dotted lines represent the variable intervals in which foils were presented.

occurring in word 1 (row 1, either /a/ or /ʌ/) and 17% for the two vowels occurring in word 2 (row 4, /æ/ or /ɛ/). In both cases, the increase over base line was sustained through all three utterances. When hearing was restored, the drop in vowel duration from the base utterances to the first utterance postswitch was 15% for the two vowels occurring in the second word (row 11). For those in the first word, there was no change detected in utterance 1 (row 7) but a drop of approximately 7% was found on the following two utterances. (Because the planned contrasts measure within-subject changes in speech parameters, the corresponding *F* values can be significant while means computed by averaging across subjects can show little or no change.)

Changes in hearing state can affect phonemic contrasts and postural variables during an ongoing segment only if the duration of that segment is long enough. Table IV shows that syllable duration means based on 30 determinations (15 trials × 2 vowels) were always well in excess of the 100–150 ms duration cited earlier in studies of *F0* compensation responses. In individual vowel utterances prior to blocking hearing (6 Ss × 15 trials × 3 utterances) not a single one had duration less than 150 ms.

2. SPL

Figure 6 and Table V summarize the results of the analysis of vowel SPL. Changes in hearing state were associated

TABLE II. Vowel contrast distances when prosthetic hearing is blocked and restored for the contrasts/a-ʌ/ and /æ-e/ (df=6,84) All contrasts are significant at  $p < 0.01$  except where *ns* (not significant) is noted.

Prosthetic hearing state	Utterance contrast	Vowel contrast	Base mean (mels)	Utterance mean (mels)	Contrast <i>F</i>	Eta <sup>2</sup> × 100	Row
Blocked	1	a-ʌ	122	124	0.7 ns	5	1
	2	a-ʌ		124	5.8	29	2
	3	a-ʌ		128	1.7 ns	11	3
	1	æ-e	90	82	7.5	35	4
	2	æ-e		91	2.8 ns	16	5
	3	æ-e		91	4.2	23	6
Restored	1	a-ʌ	128	130	1.0 ns	7	7
	2	a-ʌ		135	3.3	19	8
	3	a-ʌ		125	4.4	24	9
	1	æ-e	86	91	2.0 ns	12	10
	2	æ-e		99	7.5	35	11
	3	æ-e		93	2.7 ns	16	12

with small reliable changes in vowel SPL. All average changes were less than 1 dB. The changes in vowel SPL following blocking or restoring prosthetic hearing were all statistically significant at  $p < 0.01$  in each of the three postswitch utterances and both component words with two exceptions: viz., there was no statistically significant change from base line to the first postswitch utterance for the vowels in word 1, both when hearing was blocked (row 1) and when it was restored (row 7). In both cases, the significant change in the second utterance was sustained in the third. For the vowels in word 2 there was a small but significant sustained drop on the first utterance with hearing blocked (rows 4–6). Although one might infer that an additional delay beyond the first word of the first utterance was required for SPL changes to be expressed, the significant drop in SPL on the second word of the second utterance with hearing restored (row 11) was not sustained.

The pattern of results is not consistent with the Lombard effect, which is, however, normally measured over longer time intervals and with masking noise to block hearing. In the present study, for the most part, SPL fell when hearing was blocked and rose when it was restored, whereas in the Lombard effect the opposite occurs—louder speech with hearing blocked by masking noise and softer when only ambient noise is present (Perkell *et al.*, 2007). Some subjects showed significant changes in SPL with changes in hearing state, but others did not. The pattern of the Lombard effect was observed for only one subject, MO. Three speakers produced SPL changes contrary to the Lombard effect for hearing blocked or hearing restored. Two of these three subjects had the greatest magnitudes of change in SPL among the six subjects; thus, their response patterns were clearly reflected in the group results.

### 3. F0

Changes in hearing state were associated with small reliable changes in vowel F0 (see Fig. 7). All changes were 6% or less. The F0 changes tended to correspond to those in SPL: lower F0 when hearing was blocked, higher when it was restored. As with SPL, there was no statistically signifi-

cant change from base line to the first postswitch utterance for the vowels in word 1, both when hearing was blocked (row 1) and when it was restored (row 7). Consistent with the group results for SPL, the direction of change in fundamental frequency was opposite to that expected, that is, F0 was lowered when hearing was blocked and raised when hearing was restored. Comparing Tables V and VI, of the ten cases in which there was a change in F0 or SPL from base line to the first, second or third utterance, with either word 1 or word 2, nine out of the ten changes matched in their directions of change. The Pearson correlation coefficient between SPL and F0 for all vowels in both hearing conditions in all subjects and trials ( $n = 1440$ ) was 0.54,  $p < 0.01$ .

## IV. DISCUSSION

### A. Contrast distance

The goal of this study was to establish the time course over which adjustments in speech output are made in response to an abrupt, unanticipated change in hearing state. The responses we observed can be classified as one of two types: transient or sustained. Changes in contrast distances among vowels and sibilants were of the transient type; they did not persist reliably beyond a single utterance following a switch. Since changes in contrast distance were not reliably sustained, it appears that those changes do not function to offset or compensate for the changes in hearing introduced experimentally. Recall that, on the contrary, several studies cited in the introduction found compensatory adjustments to auditory feedback modifications. The difference may lie in this: systematic changes with short latency are observed in segmental parameters when feedback is modified, but not when it is completely blocked or restored as in the present study.

A hypothetical explanation for the differences observed with the two types of paradigm may be found in the operation of a model of speech motor planning that consists of feedforward and feedback control subsystems (cf. Guenther *et al.*, 2006). In such a model, the articulatory movements of mature speech are normally controlled by feedforward com-

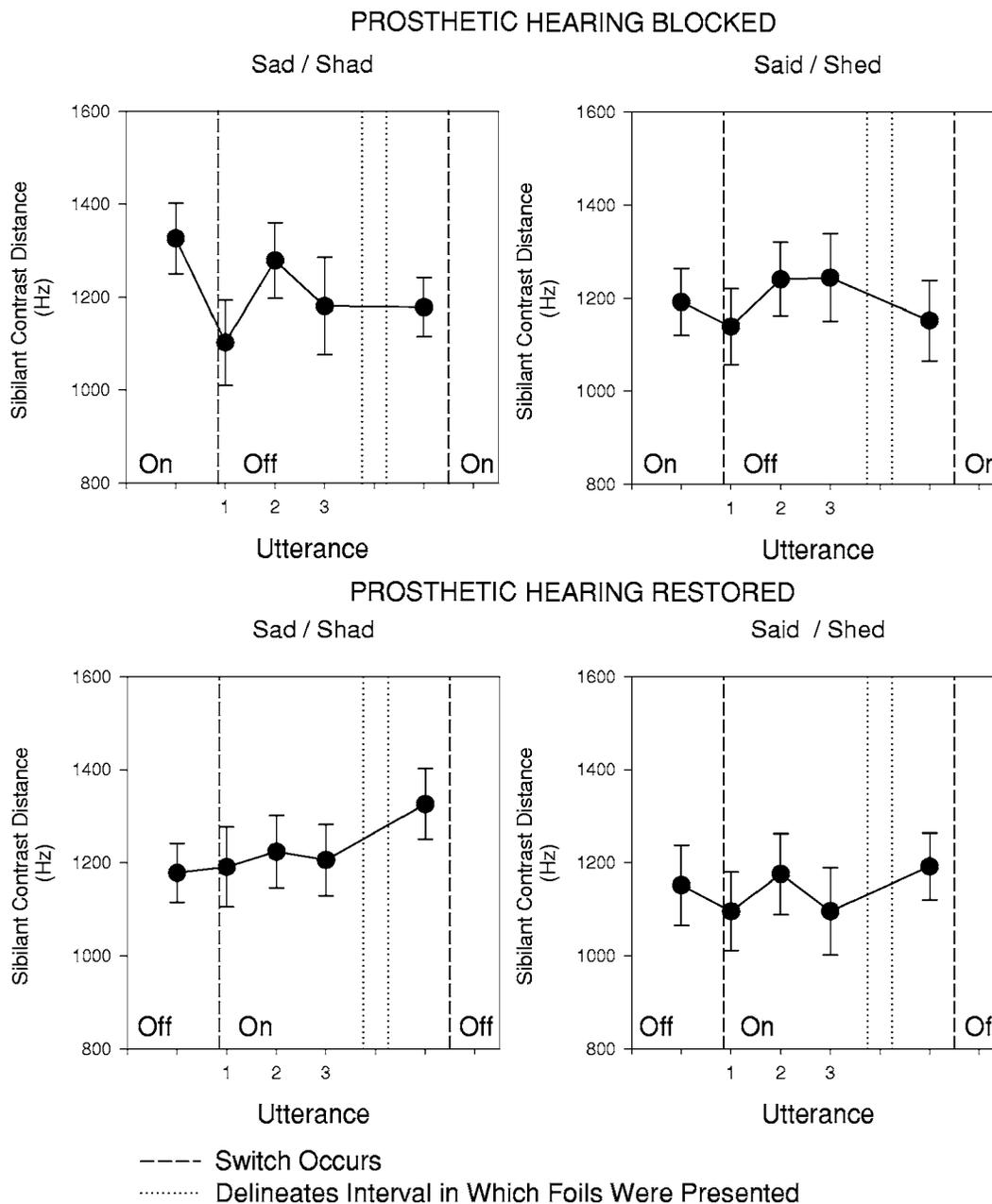


FIG. 4. The effects on sibilant contrast distance of blocking (upper panels) and restoring (lower) prosthetic hearing, as a function of utterance relative to a switch in hearing state. For details, see the caption for Fig. 2.

mands. The feedback control subsystem monitors the speech output for differences between the expected auditory (and somatosensory) consequences of producing the movements and the actual sensory results; if a large enough difference between expected and actual sensations is detected, a feedback-based corrective command is generated and, if the movement lasts long enough, the corrective motor command is expressed during that movement. As described in the introduction, such compensatory corrections have been observed in  $F_0$  (cf. Burnett *et al.*, 1997, 1998) and vowel formants (cf. Tourville *et al.*, 2005). The corrective motor commands also serve to update the feedforward motor commands for subsequent sounds (and syllables), as has been observed in the cited sensorimotor adaptation experiments (cf. Houde and Jordan, 1998, 2002; Villacorta *et al.*, 2004, 2005; Purcell and Munhall, 2006b).

Based on this view, when auditory feedback is blocked completely, there can be no feedback-based corrective motor commands; consequently, there is little or no change in produced segmental contrasts. There is also no updating of feedforward commands when feedback is blocked. Since feedforward commands are well ingrained and extremely robust, there is virtually no drift and thus the sound output remains “on target” during the time feedback remains blocked. In addition, the lack of drift in the feedforward commands makes it unlikely that any somatosensory feedback errors will be generated when hearing is blocked. When feedback is next restored, there are very few differences between auditory goals and the sound output, so again, very few segmental contrast changes are observed.

TABLE III. Sibilant contrast distances when prosthetic hearing is blocked and restored in the vowel environments /æ/ and /ɛ/. All contrasts are nonsignificant at  $p < 0.01$  ( $df=4,24$ ).

Prosthetic hearing state	Utterance contrast	Vowel context	Base mean (Hz)	Utterance mean (Hz)	Contrast $F$ All ns	Eta <sup>2</sup> × 100	Row
Blocked	1	æ	1326	1103	3.4	36	1
	2	æ		1279	0.9	13	2
	3	æ		1181	1.9	24	3
	1	ɛ	1192	1139	0.9	13	4
	2	ɛ		1240	0.4	6	5
	3	ɛ		1244	0.4	6	6
Restored	1	æ	1178	1191	0.5	7	7
	2	æ		1224	2.0	25	8
	3	æ		1205	2.8	32	9
	1	ɛ	1151	1095	2.9	33	10
	2	ɛ		1176	0.5	8	11
	3	ɛ		1096	0.5	8	12

## B. Postural parameters

Measures of postural parameters revealed some changes that were sustained over multiple utterances and persisted through the following pre-switch utterances (i.e., extreme right-hand data points in Figs. 5–7).

### 1. Vowel duration

The postural variable that changed most consistently in response to a switch in hearing state, both within and between subjects, was duration. For all four of the vowels tested (i.e., both vowels in each of the word positions 1 and 2), duration increased when hearing was blocked in the first word (*Don* or *Dun*) for all three postswitch utterances (see

Fig. 5). This made duration the only variable that changed systematically in the first word following the on-to-off switch and sustained the direction of change over all three utterances. The mean duration of the vowels in the three utterances prior to the switch that blocked hearing was greater than 150 ms, and as noted above, compensatory responses can be observed as early as about 100 ms following the introduction of an acoustic perturbation (e.g., Burnett *et al.* 1997, 1998). Comparing durations of the pre-switch vowels in this study to the shortest latencies of compensatory responses in other studies indicates that the prolongation of vowels when feedback was blocked in this study could be due to a closed-loop feedback mechanism.

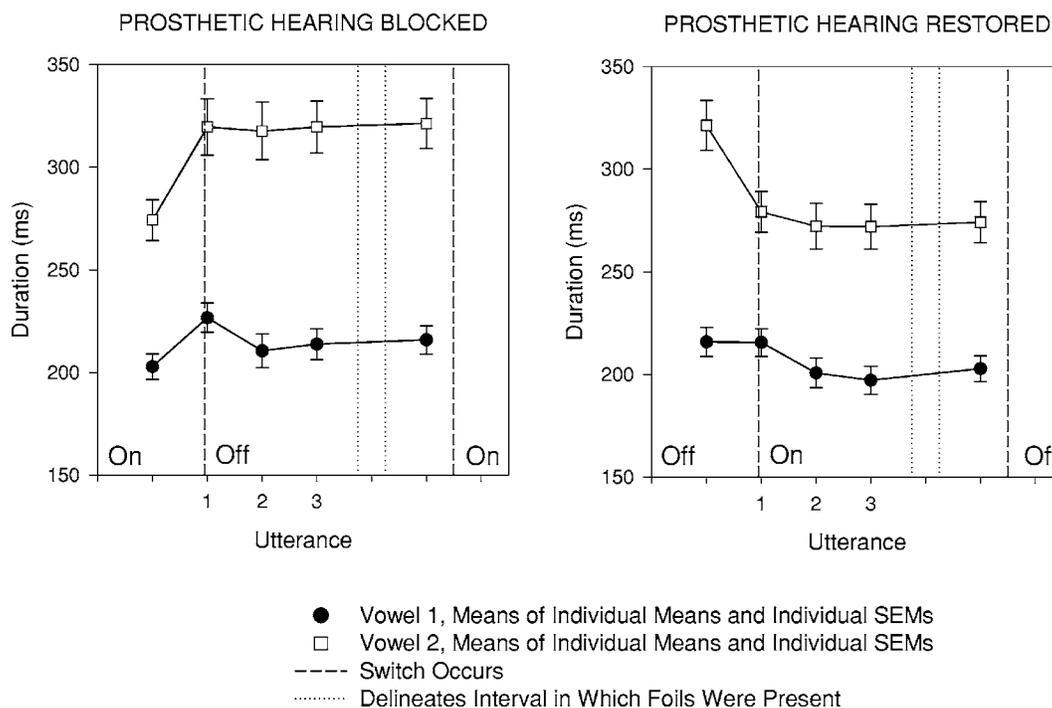


FIG. 5. The effects on vowel duration of blocking (left panel) and restoring (right panel) prosthetic hearing, as a function of utterance relative to a switch in hearing state. Vowels in the first word are represented by round symbols; vowels in the second word are represented by square symbols. Within an utterance, the vowel /a/ was always followed by the vowel /æ/; the symbols for these vowels are both filled; the vowel /a/ was always followed by the vowel /ɛ/; the symbols for these vowels are both unfilled. For further details, see caption for Fig. 2.

TABLE IV. Changes in vowel duration when prosthetic hearing is blocked and restored. All contrasts are significant at  $p < 0.01$  ( $df=6, 84$ ).

Prosthetic hearing state	Utterance positions re: switch	Vowel positions re: utterance	Base mean (ms)	Utterance mean (ms)	Contrast $F$	Eta <sup>2</sup> ×100	Row
Blocked	1	1	203	227	30	68	1
	2	1		211	5	25	2
	3	1		214	8	36	3
	1	2	274	320	54	79	4
	2	2		318	58	81	5
	3	2		320	71	83	6
Restored	1	1	216	216	4	22	7
	2	1		201	22	61	8
	3	1		197	8	36	9
	1	2	321	279	78	85	10
	2	2		272	117	89	11
	3	2		272	125	90	12

Prolongation of a sound presumably involves a combination of delaying the onset of the motor commands for producing the next sound and sustaining the agonist commands involved in producing the current one. Regarding truncation, while it might be possible to initiate motor commands for an upcoming sound earlier than planned, it should be more difficult to truncate the effects of motor commands already issued because of delays due to neural conduction and muscle contraction mechanisms. As described in Guenther *et al.* (2006), the time it takes for an action potential in a motor cortical cell to affect the length of a muscle via a subcortical motoneuron consists of “(1) the delay between motor cortex activation and activation of a muscle as measured by EMG [electromyography], and (2) the delay between EMG onset and muscle length change.” These two delays add up to

about 40 ms (Guenther *et al.*, 2006, p. 284). Presumably, additional tens of milliseconds would be required for there to be enough change in muscle length to produce a perceptible change in the acoustic output. These delays could account for the observation that vowel duration did not decrease until Vowel 2, utterance 1 when hearing was restored (Fig. 5).

## 2. Vowel SPL and $F_0$

Although there was considerable variability in the magnitude and direction of changes, if any, in SPL and  $F_0$ , they were significantly correlated so they are both discussed in this section. A surprising result was the failure to observe an appreciable increase in vowel SPL when hearing was blocked (with one exception, subject MO), and an appre-

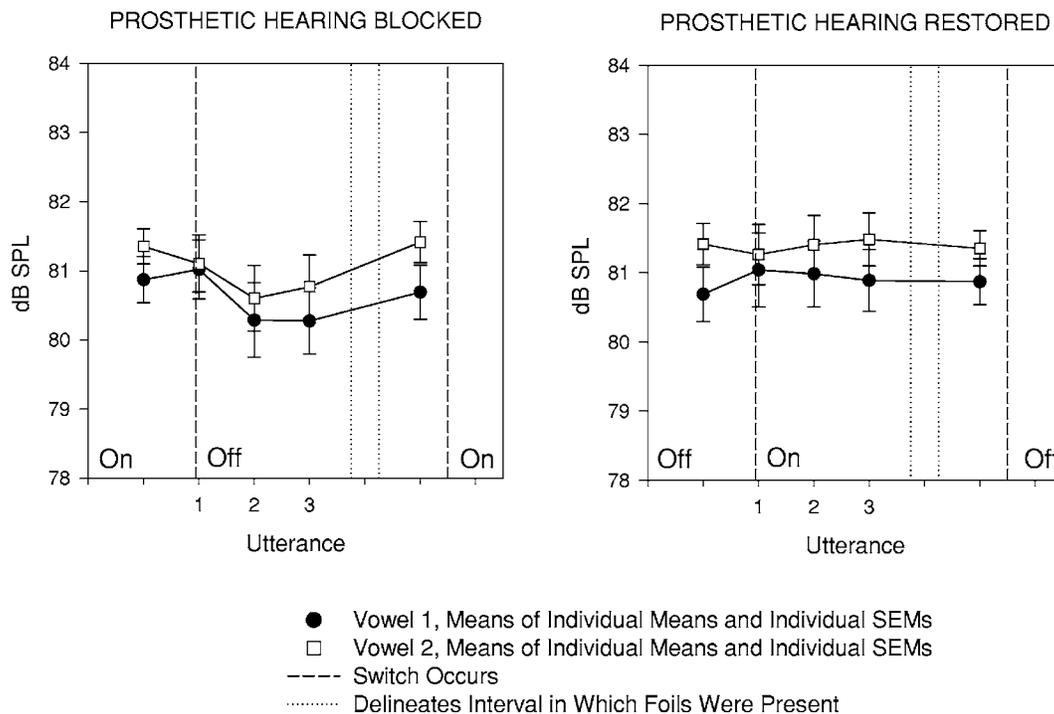


FIG. 6. The effects on vowel SPL of blocking (left panel) and restoring (right panel) prosthetic hearing, as a function of utterance relative to a switch in hearing state. For further details, see captions for Figs. 2 and 5.

TABLE V. Changes in vowel SPL when prosthetic hearing is blocked and restored. All contrasts are significant at  $p < 0.01$  except where noted; ( $df=6,84$ ).

Prosthetic hearing state	Utterance positions re: switch	Vowel positions re: utterance	Base mean (dB SPL)	Utterance mean (dB SPL)	Contrast $F$	Eta <sup>2</sup>	Row
Blocked	1	1	80.87	81.02	0.6 ns	4	1
	2	1		80.29	5	25	2
	3	1		80.27	10	41	3
	1	2		81.10	34	71	4
	2	2		80.60	24	64	5
	3	2		80.77	23	62	6
Restored	1	1	80.69	81.04	0.9 ns	6	7
	2	1		81.00	13	48	8
	3	1		80.89	15	52	9
	1	2		81.26	38	73	10
	2	2		81.40	39	74	11
	3	2		81.48	38	73	12

ciable decrease when it was restored. Two subjects responded by decreasing vowel SPL when hearing was blocked; these decreases occurred on the second utterance following the switch for vowels in both the first and second words. Three other subjects did not change vowel SPL in either direction when their speech processors were switched on or off.

When implant users' prosthetic hearing was altered in other studies from our laboratory, increased SPL was observed when prosthetic hearing was blocked and decreased when it was restored (Perkell *et al.*, 2001; Svirsky *et al.*, 1992). In those studies, however, the intervals in which auditory feedback was removed and restored—on the order of about 20 min—were far longer than the intervals of less than 1 min during the present study. We speculate that the dispar-

ity in outcomes may be attributable to the disparity in the durations of the change in hearing state. Such a conclusion would be compatible with the findings that the effects of changing auditory feedback of SPL are larger the more the task simulates real communication (Lane and Tranel, 1971). Then, too, some of our implant users may have learned, during their year's experience with their implant, to moderate their vocal level even when auditory feedback became unavailable.

The absence of auditory information is qualitatively different from the presence of loud masking noise; however, both have the effect of reducing the signal-to-noise (S/N) ratio and thus qualify as adverse speaking conditions. Lane and Tranel (1971) present evidence that speakers communicating in noise maximally increase their SPL about 5 dB for

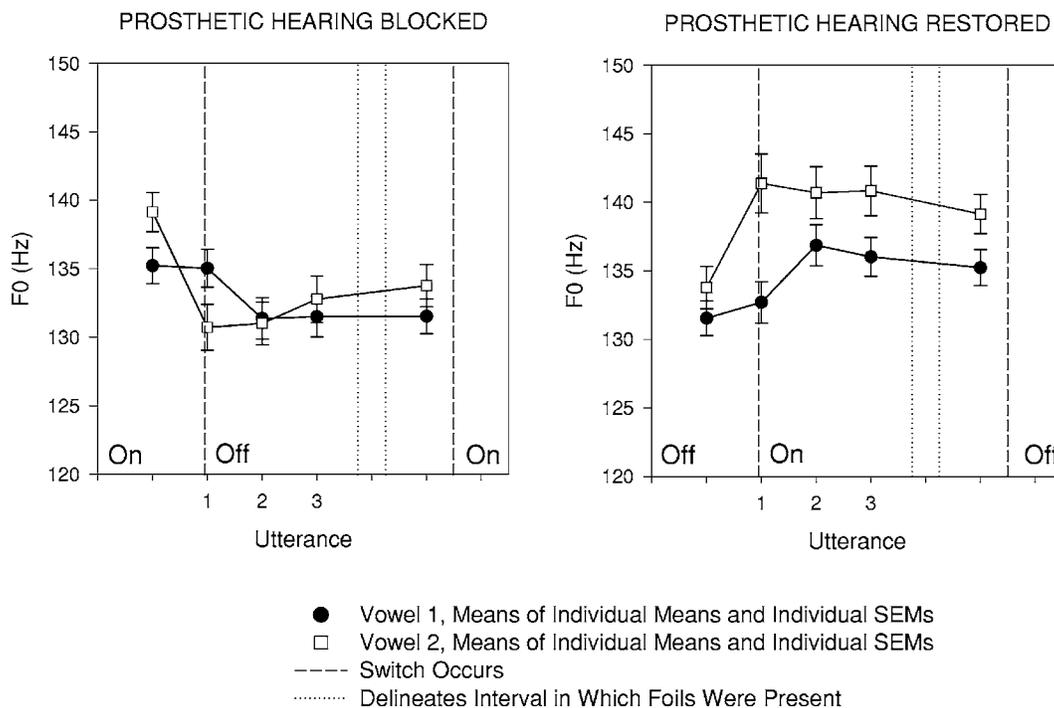


FIG. 7. The effects on vowel  $F_0$  of blocking (left panel) and restoring (right panel) prosthetic hearing, as a function of utterance relative to a switch in hearing state. For further details, see captions for Figs. 2 and 5.

TABLE VI. Changes in vowel fundamental frequency when prosthetic hearing is blocked and restored. All contrasts are significant at  $p < 0.01$  except where nonsignificant (ns) is noted. ( $df = 6, 84$ ).

Prosthetic hearing state	Utterance positions re: switch	Vowel positions re: utterance	Base mean (Hz)	Utterance mean (Hz)	Contrast $F$	Eta <sup>2</sup> × 100	Row
Blocked	1	1	135	135	3 ns	15	1
	2	1		132	54	80	2
	3	1		132	155	92	3
	1	2	139	131	72	84	4
	2	2		131	80	85	5
	3	2		133	70	83	6
Restored	1	1	132	133	2 ns	12	7
	2	1		137	149	91	8
	3	1		136	81	85	9
	1	2	134	141	20	59	10
	2	2		141	70	83	11
	3	2		141	108	88	12

every increase in noise of 10 dB and they term this relation the “Lombard function.” It has been suggested that the tendency of the deaf to speak more loudly is likewise attributable to compensation for the apparent reduction of the S/N ratio. An experiment by Black (1951) supports this interpretation. Black had 144 college students read lists of words while their SPL was recorded. Then he reduced speakers’ self-hearing by exposing them to 110 dB noise for 2 h. After the termination of the noise, he measured subjects’ temporary hearing loss at 3 min intervals and again recorded their vocal level several times. Lane and Tranel (1971) report that the function relating the speaker’s vocal level in Black’s experiment to the amount of that speaker’s temporary hearing loss had approximately the same slope (in dB coordinates) but opposite sign as the Lombard function. Lane, Tranel and Sisson (1970) present other evidence that this “sidetone compensation function” is approximately the reciprocal of the Lombard function. Although these studies were conducted with normal-hearing subjects, Perkell *et al.* (2007) found functions relating vocal level to signal-to-noise ratio that were somewhat similar for a group of speakers with normal hearing and one with cochlear implants, when subjected to masking of auditory feedback.

Speakers with normal hearing can learn to modulate the Lombard effect. Those studied by Pick *et al.* (1989) spoke during alternating periods of quiet and noise. When given visual feedback about their vocal level and training in regulating it, they were able to speak quietly in noisy backgrounds, even after the visual feedback was removed. Subjects who returned for retesting a day later were able to restrict the Lombard effect “sharply” without visual feedback. Interestingly, some of these subjects “overcompensated” by dropping their vocal intensity in noise below the levels they used in quiet. Two of the subjects in the current study also used lower vocal intensity when auditory feedback became unavailable.

Just as speakers can voluntarily minimize changes in SPL despite changing listening conditions, so too they can produce voluntary responses to pitch-shifted auditory feedback. Hain *et al.* (2000) instructed subjects to respond to a

pitch shift either by raising or lowering their  $F_0$ , or by ignoring it and keeping  $F_0$  constant. In addition to the expected compensatory change opposite to the direction of the shift, with a latency of 100–150 ms, Hain *et al.* observed a later response, which was almost always made in accordance with experimenters’ instructions, hence in the same direction as the shift when so instructed. This provides another example of a response to a change in auditory feedback that can be modified by experimenter’s instructions.

## V. SUMMARY AND CONCLUSIONS

Returning to the hypotheses of the current study, the first part of Hypothesis 1 predicted that when hearing was blocked, segmental contrasts would decrease and they would increase when hearing was restored. This part of the hypothesis was disconfirmed. With minor exceptions, there were no reliable changes in segmental parameters. We attribute this result to the nature of the perturbation—removal and restoration of feedback, as opposed to modification of feedback parameters. According to Guenther’s DIVA model of speech motor control (Guenther *et al.*, 2006), removal and restoration of auditory feedback would not result in any feedback-based error corrections to current movements or to feedforward commands for future movements, so feedforward commands should remain unchanged in the short run, regardless of whether or not feedback is present.

The second part of Hypothesis 1 predicted that blocking auditory feedback would result in increases in SPL,  $F_0$ , and sound segment durations; restoration of hearing would reverse these changes. Such predictions are compatible with earlier results on the behavior of these speech postural parameters in response to changes in acoustic transmission conditions (cf. Lane and Tranel, 1971; Perkell *et al.*, 1992, 2007; Svirsky *et al.* 1992). In support of this prediction, the current results showed increases in sound segment durations when feedback was blocked and decreases when hearing was restored; the duration changes were sustained until the next switch in hearing state. SPL and  $F_0$  values were correlated, most likely due (at least partly) to the aerodynamic/

biomechanical interdependence of these two parameters (cf. Perkell *et al.*, 1992). In contrast to the findings on duration, changes in SPL were not systematic. They differed among subjects; some were opposite to the direction found in other studies; they were only observed on the first utterance after blocking hearing for one of the two vowel pairs. These anomalous findings may be due to the brevity of the changes in processor state that were employed or to the tendency of some subjects, when they do not have access to prosthetic hearing from their implants, to actively suppress increases in SPL in order to avoid adverse social consequences of speaking too loudly.

Hypothesis 2 predicted that parameter changes would take place in the word following the one in which the switch was made, unless the duration of the vowel in the first word exceeded 150 ms, in which case, responses could occur during that vowel. This hypothesis was supported for duration measurements (but not SPL and *F0*) when hearing was blocked and also when hearing was restored. With a switch in hearing state introduced unexpectedly at the beginning of the vowel in the first of the two CVC words, the duration of the vowel in the first word after the switch exceeded 150 ms, and indeed the first and subsequent vowels were altered in duration as soon as possible after the switch—lengthened during Vowel 1, Utterance 1 with hearing blocked and shortened during Vowel 2, Utterance 1 with hearing restored.

Finally, Hypothesis 3 predicted differences in the latency of parameter changes following a switch in hearing state, depending on whether the parameter indexes a segmental or postural speech variable. As it turned out, changes in contrast distances among vowels and sibilants were of the transient type; they did not persist beyond a single utterance following a switch, whereas duration, a postural variable, changed in the first word following the change in hearing state and sustained that change.

The production of *sound segments*, syllables and words under feedforward control very rarely encounters the kind of distorting perturbations that would cause mismatches between auditory goals and the produced sounds. On the other hand, changes in acoustic transmission conditions, such as in the level of environmental noise, occur frequently and elicit short-latency responses in the form of changes in speaking rate and level. As presented above, the results for segmental and postural parameters did differ from one another. Because unexpectedly blocking and restoring auditory feedback cannot engage the postulated feedback control system that is involved in segmental (and syllabic) sound production, no systematic changes in vowel and sibilant contrasts were observed. The contrary finding of some systematic postural changes that appear to be implemented by feedback control leads to the inference that the control of segmental and postural aspects of speech may involve somewhat different mechanisms, as proposed previously by Perkell *et al.* (1992).

## 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 Infirmary 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.

- 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.
- Burnett, T. A., Senner, J. E., and Larson, C. R. (1997). "Voice *F0* responses to pitch-shifted auditory feedback: A preliminary study," *J. Voice* **11**, 202–211.
- Burnett, T. A., Freedland, M. B., Larson, C. R., and Hain, T. C. (1998). "Voice *F0* responses to manipulations in pitch feedback," *J. Acoust. Soc. Am.* **103**, 3153–3161.
- Clark, J. E., Lubker, J. F., and Hunnicutt, S. (1987). "Some preliminary evidence for phonetic adjustment strategies in communication difficulty," in *Language Topics: Essays in Honor of Michael Halliday*, R. Steele and T. Threadgold, eds., pp. 161–180 (Benjamins, Amsterdam).
- Cowie, R. I. D., and Douglas-Cowie, E. (1983). "Speech production in profound postlingual deafness," in *Hearing Science and Hearing Disorders*, M. Lutman and M. P. Haggard, eds., pp. 183–230. (Academic, London).
- Draeger, 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.
- 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.
- 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.
- Hain, T. C., Burnett, T. A., Kiran, S., Larson, C. R., Singh, S., and Kenney, M. K. (2000). "Instructing subjects to make a voluntary response reveals the presence of two components to the audio-vocal reflex," *Exp. Brain Res.* **130**, 133–141.
- 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.
- Houde, J. F., and Jordan, M. I. (1998). "Sensorimotor adaptation in speech production," *Science* **279**, 1213–1216.
- Houde, J. F., and Jordan, M. I. (2002). "Sensorimotor adaptation of speech I: Compensation and adaptation," *J. Speech Lang. Hear. Res.* **45**, 295–310.
- Jones, J. J., and Munhall, K. G. (2002). "The role of auditory feedback during phonation: Studies of Mandarin tone production," *J. Phonetics* **30**, 303–320.
- Jongman, A., Wayland, R., and Wong, S. (2000). "Acoustic characteristics of English fricatives," *J. Acoust. Soc. Am.* **108**, 1252–1263.
- Kawahara, H., and Williams, J. C. (1996). "Effects of auditory feedback on voice pitch trajectories: Characteristic responses to pitch perturbations," in *Vocal Fold Physiology*, P. J. Davis and N. H. Fletcher, eds., pp. 263–278 (Singular, San Diego).
- 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.
- Lane, H., Denny, M., Guenther, F. H., Matthies, M., Ménard, L., Perkell, J., 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., Denny, M., Guenther, F., Perkell, J., Stockmann, E., Tiede, M., Vick, J., and Zandipour, M. (in press, 2007). "Effects of short-

- and long-term changes in auditory feedback on vowel and sibilant contrasts," J. Speech, Lang. Hear. Res.
- Lane, H., Matthies, M., Perkell, J., Vick, J., and Zandipour, M. (2001). "The effects of changes in hearing status in cochlear implant users on the acoustic vowel space and CV coarticulation," J. Speech Lang. Hear. Res. **44**, 552–563.
- Lane, H., Tranel, B., and Sisson, C. (1970). "Regulation of voice communication by sensory dynamics," J. Acoust. Soc. Am. **47**, 618–624.
- Lane, H., and Tranel, B. (1971). "The Lombard sign and the role of hearing in speech," J. Speech Hear. Res. **14**, 677–709.
- Lane, H., and Webster, J. W. (1991). "Speech deterioration in postlingually deafened adults," J. Acoust. Soc. Am. **89**, 859–866.
- Lindblom, B. E. F. (1990). "Explaining phonetic variation: A sketch of the H&H theory," in *Speech Production and Speech Modeling*, pp. 403–439 (Kluwer, Dordrecht).
- Markel, J. D., and Gray, A. H. (1976). *Linear Prediction of Speech* (Springer-Verlag, Berlin).
- 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., Perkell, J., 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**, 936–946.
- McKay, C. M., and McDermott, H. J. (1993). "Perceptual performance of subjects with cochlear implants using the Spectral Maxima Sound Processor (SMSPP) and the Mini Speech Processor (MSP)," Ear Hear. **14**, 350–367.
- Natke, U., and Kalveram, K. T. (2001). "Effects of frequency-shifted auditory feedback on fundamental frequency of long stressed and unstressed syllables," J. Speech Lang. Hear. Res. **44**, 577–584.
- Perkell, J. S., Denny, M., Lane, H., Guenther, F. H., Matthies, M. L., Tiede, M., Vick, J., Zandipour, M., and Burton, E. (2007). "Effects of masking noise on vowel and sibilant contrasts in normal-hearing speakers and postlingually deafened cochlear implant users," J. Acoust. Soc. Am. **121**, 505–514.
- Perkell, J., Lane, H., Svirsky, M., and Webster, J. (1992). "Speech of cochlear implant patients: A longitudinal study of vowel production," J. Acoust. Soc. Am. **91**, 2961–2978.
- 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.
- Peters, R. W. (1955). "The effect of filtering of sidetone on speaker intelligibility," J. Speech Hear. Disord. **20**, 371–375.
- Pick, H. L., Jr., Siegel, G. M., Fox, P. W., Garber, S. R., and Kearney, J. K. (1989). "Inhibiting the Lombard effect," J. Acoust. Soc. Am. **85**, 894–900.
- Purcell, D. W., and Munhall, K. G. (2006a). "Compensation following real-time manipulation of formants in isolated vowels," J. Acoust. Soc. Am. **119**, 2288–2297.
- Purcell, D. W. and Munhall, K. G. (2006b). "Adaptive control of vowel formant frequency: Evidence from real-time formant manipulation," J. Acoust. Soc. Am. **120**, 966–977.
- Stevens, K. N., Nickerson, R. S., and Rollins, A. M. (1983). "Suprasegmental and postural aspects of speech production and their effect on articulatory skills and intelligibility," in *Speech of the Hearing-Impaired*, pp. 35–51 (University Park Press, Baltimore).
- Svirsky, M. A., Lane, H., Perkell, J. S., and Wozniak, J. (1992). "Effects of short-term auditory deprivation on speech production in adult cochlear implant users," J. Acoust. Soc. Am. **92**, 1284–1300.
- 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.
- Tourville, J. A., Guenther, F. H., Ghosh, S. S., Reilly, K. J., Bohland, J. W., and Nieto-Castanon, A. (2005). "Effects of acoustic and articulatory perturbation on cortical activity during speech production," in *11th Annual Meeting of the Organization for Human Brain Mapping*, Toronto, June 12–16, p. S49.
- 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.
- Villacorta, V., Perkell, J. S., and Guenther, F. H. (2004). "Sensorimotor adaptation to acoustic perturbations in vowel formants," J. Acoust. Soc. Am. **115**, 2430(A).
- Villacorta, V., Perkell, J. S., and Guenther, F. H. (2005). "Relations between speech sensorimotor adaptation and perceptual acuity," J. Acoust. Soc. Am. **117**, 2618–2619(A).
- Waldstein, R. S. (1990). "Effects of postlingual deafness on speech production: Implications for the role of auditory feedback," J. Acoust. Soc. Am. **88**, 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.
- Xu, Y., Larson, C. R., Bauer, J. J., and Hain, T. C. (2004). "Compensation for pitch-shifted auditory feedback during the production of Mandarin tone sequences," J. Acoust. Soc. Am. **116**, 1168–1178.