

December 01, 1995

Phonemic resetting versus postural adjustments in the speech of cochlear implant users: an exploration of voice-onset time

Harlan Lane

Northeastern University; Massachusetts Institute of Technology

Jane Wozniak

Massachusetts Institute of Technology

Melanie Matthies

Massachusetts Institute of Technology

Mario Svirsky

Massachusetts Institute of Technology

Joseph Perkell

Massachusetts Institute of Technology

Recommended Citation

Lane, Harlan; Wozniak, Jane; Matthies, Melanie; Svirsky, Mario; and Perkell, Joseph, "Phonemic resetting versus postural adjustments in the speech of cochlear implant users: an exploration of voice-onset time" (1995). *Psychology Faculty Publications*. Paper 14.

<http://hdl.handle.net/2047/d20000861>

Phonemic resetting versus postural adjustments in the speech of cochlear implant users: An exploration of voice-onset time

Harlan Lane,^{a),b)} Jane Wozniak,^{a)} Melanie Matthies, Mario Svirsky, and Joseph Perkell
*Massachusetts Institute of Technology, Research Laboratory of Electronics, Room 36-511,
50 Vassar Street, Cambridge, Massachusetts 02139*

(Received 4 November 1994; accepted for publication 15 June 1995)

Voice-onset time (VOT) was measured in plosive-initial syllables uttered by five cochlear implant users prior to and repeatedly at intervals after activation of their speech processors. In “short-term” experiments, the elicitation set was read after the subject’s processor had been off for 24 h, then turned on, then off again. Four out of five implant users increased voiceless and/or voiced VOTc (VOT corrected for changes in syllable duration) from preimplant baselines to final recordings made 1–3 years later. Measured acoustic correlates of speech “posture” (average SPL, F_0 , and low-frequency spectral slope) changed concurrently. Results in the short-term study were largely consistent with the long term. Significant multiple regressions relating changes in VOTc to accompanying changes in postural correlates were found in both studies. This outcome is consistent with hypotheses that predict changes in both VOTc and in postural correlates with the restoration of some hearing and that allow for linkages between the two. Some of the reliable VOTc increases obtained over the long term that were not correlated with postural changes may have been caused directly by auditory validation of articulatory/acoustic relations that underlie synergisms for phoneme production. © 1995 Acoustical Society of America.

PACS numbers: 43.70.Dn, 43.70.Bk, 43.70.Fq, 43.71.Ky

BACKGROUND

In a study of vowel production that examined four speakers before and for up to 2 years after they received Ineraid cochlear implants, Perkell *et al.* (1992) reported numerous changes in parameter values with the restoration of some hearing, including changes in formants, voice sound-pressure level, fundamental frequency, duration, an indirect measure of breathiness, and airflow; many of these changes were correlated. Some of the changes may have been the direct result of the speaker hearing a given phonemic speech parameter for the first time in many years and making an articulatory adjustment to bring anomalous parameter values into line with phonemic intentions. Other changes, like those in vowel SPL, may have been the result of the speaker recovering some ability to monitor conditions for the transmission of spoken messages and making a “postural” adjustment to ensure a certain level of intelligibility. A posture is a state of the production mechanism inferred from an average parameter value, on which meaningful fluctuations are imposed. Examples of speech postures are the balance between expiratory and inspiratory forces associated with a subglottal pressure, average tension of the vocal folds, average degree of adduction of the glottis, average position of the tongue body, and speaking rate. These postures affect the average values of SPL, F_0 , $H1-H2$ (amplitude difference between

the first two harmonic peaks in the log magnitude spectrum, a measure of low-frequency spectral slope), $F1$ and $F2$, and syllable duration.

Although the two roles we have postulated for auditory feedback are distinct, their consequences frequently are not because phonemic and postural speech parameters are often physically interdependent—aerodynamically, acoustically, and mechanically. For example, speakers in this vowel study increased their speaking rate after receiving prosthetic hearing (also see Leder *et al.*, 1987a). As mean vowel duration fell across recording sessions, mean first-formant frequency also tended to fall, especially for low vowels. The investigators proposed that the $F1$ changes were a consequence of rate changes (at higher rates there is less time for jaw and tongue lowering for the low vowels, so average tongue height increases) and they point out that, since those changes compress the $F1$ range, they are unlikely to have been made actively in response to perceived $F1$ abnormality.

If research on the speech of deafened adults who receive prostheses is to throw light on the role of hearing in regulating speech, it will be necessary to distinguish changes in implementing phonemic contrasts from changes in posture. To put the problem in context, we start with the perspective that articulatory movements are programmed to achieve sequences of feature-specified goals and that this programming uses an acquired internal model of relations between articulatory commands and acoustic output (Perkell *et al.*, 1995). We hypothesize that auditory feedback is used to validate the articulatory/acoustic relations specified in the internal model. Because speech transmission takes place under variable conditions (for example, ambient noise), audition has a second

^{a)}Also of Department of Otolaryngology, Massachusetts Eye and Ear Infirmary, Harvard Medical School, Boston, MA 02114.

^{b)}Also of Northeastern University, Boston, MA 02115.

and more immediate role in assuring the production of intelligible messages. The speaker responds to hearing such changing transmission conditions with changes in postural settings, in order to ensure that speech is loud enough and slow enough to be understood.

When some hearing is restored with a cochlear implant, the deafened speaker will respond promptly to the changed transmission conditions and will make postural adjustments. Psychophysical findings with implant users indicate that auditory monitoring of transmission conditions and regulation of posture may be facilitated because many of the acoustic correlates of posture, such as average SPL or F_0 , are quite audible to implant users (Shannon, 1993). The implant user may also detect the discrepancy, for a given speech sound, between phonemic intentions and the acoustic results (this depends in part on the properties of the individual's prosthetic "audition"). If so, he or she may revise the articulatory parameter values for implementing that phoneme. The effects of changing these phonemic settings are frequently confounded in acoustic-phonetic measures with the effects of postural adjustments because the same articulators may be involved in both (e.g., position of the tongue body), or because of the many physical linkages in speech.

In order to examine changes in postural and phonemic settings brought about by changes in hearing, and the ways in which measures of each may be confounded, this study focuses on the voicing contrast in English. Among the several acoustic correlates of the voicing contrast, voice-onset time has been studied most extensively as a perceptual cue and as a parameter of phoneme production in both hearing-impaired and normally hearing subjects. Lane *et al.* (1994) have reported VOT measurements obtained from four deafened speakers who characteristically uttered plosives with too-short voice-onset time. After extended use of cochlear prostheses, three of the four increased voiced and/or voiceless VOT (corrected for changes in syllable duration). Because the voiceless increases were larger than the voiced, the difference between contrasting phonemes was enhanced. These speakers were also able to identify plosives quite accurately using their prostheses on listening tests, a gross indication that they could have been guiding their VOT changes using their prosthetic audition.

There are many postural adjustments, starting with speaking rate, that could be responsible for, or contribute to, the observed changes in VOT following activation of the subjects' speech processors, thus calling into question the view that it was the subjects' newfound ability to hear their voice-onset times that led them to modify VOT. In articulatory terms, VOT is the interval between release of the supraglottal occlusion and the onset of vocal fold vibration. Because voiceless plosives are characterized by an active abduction of the vocal folds (Hirose, 1976; Löfqvist, 1980; Weismer, 1984), laryngeal gesture may be the primary determinant of their VOT. Therefore, with phonetic context constant, VOT in voiceless plosives is affected by the timing of the peak glottal opening, the magnitude of that opening, the rate of glottal closure, and the duration of the following vowel. VOT in the voiced plosives, on the other hand, may be secondary to the timing of the drop in transglottal pres-

sure (cf. Sernicales *et al.*, 1984), which is due to the release of the supraglottal constriction as well as to such postures as the balance between expiratory and inspiratory forces associated with a subglottal pressure and average tension and stiffness of the vocal folds.

Although for simplicity we contrast acoustic correlates of postures with acoustic correlates of phonemes, it should be kept in mind that the latter have both a phonemic component and an overall postural component. Thus, for example, $H1-H2$, measured early in the vowel following a plosive, presumably depends not only on the average glottal posture that the subject adopts but also on whether the plosive is voiced or voiceless.

According to the analysis of the physics of laryngeal behavior in Stevens (1977), we may expect F_0 and VOT to be positively correlated, $H1-H2$ and VOT to be positively correlated, and SPL and VOT to be negatively correlated. (Several studies have shown that vowel duration and VOT are positively correlated; see below.) One link between F_0 and VOT is vocal fold stiffness, which may be achieved by contraction of the cricothyroid and vocalis muscles. The closing portion of the glottal vibratory cycle is due to the restoring force of vocal fold stiffness (and to Bernoulli forces). An increase in stiffness not only increases F_0 but also reduces the range of subglottal pressures and glottal apertures sufficient to sustain vocal fold vibration. Raising the larynx also increases longitudinal tension on the folds and can reduce the intraoral volume, inhibiting airflow through glottis and vocal fold vibration. Turning to $H1-H2$, if the arytenoid cartilages are spread, complete glottal closure never occurs and airflow is cut off for a portion of the cycle only in the anterior region of the glottis. The resulting glottal waveform has no sharp discontinuities and less energy at high frequencies, hence a steeper spectral slope and therefore higher values of $H1-H2$ (Klatt and Klatt, 1990). For a given fold stiffness and subglottal pressure, greater arytenoid spacing inhibits voicing; hence $H1-H2$ and VOT will be directly related. Finally, with higher transglottal pressure drop, there is higher SPL (Holmberg *et al.*, 1994) and a broader range over which vocal fold vibration can be sustained (hence, shorter VOT) despite spreading of the glottis, stiffening of the vocal folds, or raising of the larynx.

From this brief sketch it is apparent that there are complex interactions between posture and VOT and that the functional relations between indirect indices of posture and voicing are unlikely to be straightforward.

Not only can postural changes mediate an observed change in VOT following cochlear prosthesis, but also the inverse result—failure to find a change in speech after prosthesis—can also be misleading if posture is not taken into account. Tartter *et al.* (1989) reported that a teenager deafened at age six did not change VOT reliably after using a cochlear implant for a year. However, their subject's speaking rate increased over the year, and this reduction in syllable duration may have entrained reduced VOT since VOT is shorter in shorter syllables (Volaitis and Miller, 1992). Rate-induced reductions in VOT could offset increases in VOT from phonemic resetting. Waldstein (1990) measured VOT with seven deafened speakers and seven hearing and found

that the deaf speakers had too-short VOT for voiceless plosives. She concludes, "Auditory feedback serves to fine-tune the VOT values typical of English in adulthood" (p. 2104). However, until we investigate the possible contributions of postural changes to VOT changes, we cannot rule out the possibility that the shortened VOTs of her deafened speakers were due not to a lack of opportunity for auditory validation of VOT itself but rather to the effects of posture (or both). Lane *et al.* (1994) concluded that they could not rule out the possibility that the VOT changes they reported were brought about indirectly through the mediation of another speech production mechanism also affected by the processor activation.

Acoustic parameter values of the phonemes are determined simultaneously by posture, prosody, and the "intrinsic" properties of the phoneme itself. Following the restoration of some hearing, there will be changes in speech postures, some of which are linked to the production mechanisms governing the phoneme, changes in the prosody of utterances in which the phoneme of interest is embedded, and changes in the phonemic settings of articulatory routines underlying the production of the phoneme itself, the result of selective effects of the speaker's renewed ability to perceive the relevant phonemic contrasts—what we have called "auditory validation."

In this study, we examine the effects of processor activation on a measure derived from plosive-initial syllables that takes into account the effects of rate on VOT. The study seeks to partial out the additional sources of variance due to postural changes other than rate by regressing indices of postural changes on VOT changes. This statistical approach is necessitated by limitations in direct measurement of speech posture and by the lack of sufficiently comprehensive quantitative models for assessing interdependences. Two experimental paradigms are employed. In the longitudinal study, recordings were made of the speech of five deafened adults before and shortly after the speech processors of their cochlear implants were activated for the first time, and repeatedly thereafter over a period of several years. In "short-term" experiments, subjects turned off their processors for 24 h before coming to the laboratory, where they read an elicitation set with their processors left off, turned on, and turned off again.

The hypotheses guiding this study were (1) when some hearing is restored with cochlear prosthesis, speakers will not only increase the VOT values of voiced and voiceless stops (if they were abnormally short), but will also make many concurrent changes in posture indexed by measures such as *H1-H2*, *SPL*, and *F0*, reflecting auditory monitoring of transmission conditions. (2) As a further result of auditory monitoring, when speakers' processors are turned off for brief periods, acoustic parameter values indexing posture will regress toward their preimplant values, since transmission conditions will have reverted to those in effect during long years of deafness. When implant processors are turned on again, changing the transmission conditions, postural indices will recover recent average values, reflecting auditory monitoring. (3) With prosthetic hearing available, increases in VOT will frequently be attributable to postural changes.

Increases in VOT that cannot be linked to postural changes may reflect phonemic resetting due to auditory validation.

I. METHOD

A. Experiment I: Longitudinal modification

The rationale for the within-subjects repeated-measures design and recording and analysis procedures will be found in detail in Lane *et al.* (1994); a summary is given here with details of additional analyses.

1. Subjects

The three female subjects and one male subject were the same as in Lane *et al.* (1994), where additional subject characteristics are given. A second male subject, MD, has been added. Subject FA had a congenital monaural impairment and wore a hearing aid until she became profoundly deaf at age 33. Subject FB had normal hearing until age 21 and bilateral progressive hearing loss, partially corrected with hearing aids, until age 40, when she became profoundly deaf. Subject FC had a severe bilateral hearing loss since early childhood and wore hearing aids until age 47, when she stopped using them. Male subject MC had a progressive bilateral hearing loss beginning at age ten and wore a hearing aid from that time until approximately 6 months after the activation of the speech processor of his cochlear implant. MD had a congenital progressive binaural loss, partially corrected with hearing aids worn until the time he received a cochlear implant. All of the subjects had pure-tone average losses greater than 102 dB in each ear prior to implant. All subjects used their cochlear implants regularly. Their scores on the NU-6 test of word recognition, auditory only, ranged from 8% to 36% correct after 1–4 years of implant use.

2. Prosthesis

The Ineraid cochlear implant (Richards Medical Co.) consists of an implanted electrode array, a percutaneous pedestal and connector, and an external sound processor. The sound processor has an ear level microphone, a wideband automatic gain control, and four overlapping bandpass filters with crossover frequencies of approximately 0.7, 1.4, and 2.3 kHz. The four analog filter outputs are delivered (via the percutaneous connector) individually to four monopolar intracochlear electrodes, with a common return electrode. The electrodes, spaced 4 mm apart, were positioned successfully in all subjects by insertion into the scala tympani through the round window, with the first placed most apically, some 22 mm from the round window. Gain controls include user adjustments for input sensitivity and volume, and channel specific gains that are set for each subject.

3. Speech elicitation

Two baseline recordings of speech production were obtained from each subject before activating the speech processor of the subject's cochlear implant. Postactivation recordings were made at intervals of approximately 0, 4, 12, 26, 52, and 104 weeks after the speech processor of the implant was turned on; in addition, subject FA was recorded at 141, 210, and 260 weeks. The speech material consisted of the six

English plosives spoken in the carrier phrase “It’s a /Cad/ again.” These utterances were arranged in a quasirandom sequence read three times (five times for MD); other speech material was recited for approximately 20 min between each reading.

4. Recording, calibration, signal processing, and data analysis

The subject was seated in a sound-attenuating room. A small electret microphone was placed at a fixed distance of 20 cm in front of the subject’s lips by attaching it to a flexible arm affixed to the back of the chair. The utterance materials were projected on a screen located several feet in front of the subject. For calibration of sound-pressure level, a sound source (electrolarynx) was placed in front of the subject’s lips, while an experimenter observed the sound level value on an SPL meter (C scale) held next to the electret microphone. The microphone signal was amplified, recorded, and later low-pass filtered at 4.8 kHz and digitized at 10 kHz. Digitization, signal processing, and interactive data extraction were performed with procedures written in the MITSYN languages running on a Digital Equipment Corporation engineering workstation (Perkell *et al.*, 1991).

An experimenter, working with a display of the digitized speech signal of each utterance, placed markers at the onset of the plosive release burst (t_1), at the first zero crossing at the onset of periodicity in the waveform (t_2), and at the zero crossing following the last periodic pulse of the vowel (t_3). VOT was computed as the interval between t_1 and t_2 , and duration as the interval between t_1 and t_3 , the CV portion of the CVC syllable.

In order to assess speech postures close to the time of each VOT value, three measures were made on the vowel following the plosive in each utterance. The measures were made at a point 20 ms following the onset of the vowel, using a 51.2-ms window. The measures were (1) $H1-H2$ —the amplitude difference between the first two harmonics in the acoustic spectrum, corrected for the influence of $F1$ and with pre-emphasis of 6 dB/oct. This parameter is a measure of slope of the low-frequency region of the source spectrum and the degree to which the underlying flow waveform is sinusoidal in shape; it is related to the perceived “breathiness” of the voice and presumably the degree of glottal abduction (Holmberg *et al.*, 1988; Klatt and Klatt, 1990). $H1-H2$ was not measurable on 12 of 18 voiced tokens of FC and 11 of 18 for MC. (2) SPL—To determine the SPL of each vowel token, the rms of the recorded, digitized sound-pressure signal was expressed as a proportion of that of the calibration tone and converted to dB. The interval from 20 ms after vowel onset to 20 ms before vowel offset was delimited and the time one-fourth of the way through the interval identified. The 51.2-ms window was centered at that time and the average value of dB SPL in that window retained. (3) $F0$ —An algorithm supplied with the MITSYN languages was used to track and display individual periods of the voice fundamental. When the tracking was not optimal an operator adjusted one or more tracking parameters. The same method to position the analysis window was used as with SPL and the average value of $F0$ in the window retained.

5. Correction for speaking rate

The syllable durations of implant users characteristically shorten with processor activation (Perkell *et al.*, 1992; Leder *et al.*, 1987a). VOT, in turn, varies directly with syllable duration in both speakers with normal hearing (Summerfield, 1975; Volaitis and Miller, 1992; Diehl *et al.*, 1980; Pind, 1995) and in implant users (Lane *et al.*, 1994). Thus, even when the speaker increases phonemic settings for VOT as a result of his or her newfound ability to hear correlates of voicing, VOT following processor activation could actually decrease due to rate changes. To correct VOT measures for the effect of rate, the size of the effect was first estimated. VOT was regressed on syllable duration for data from each of the three speakers who participated in the magnitude production study of VOT and rate, reported by Volaitis and Miller (1992). (Only syllables with durations between 200 and 400 ms were retained for the regression as that was approximately the range of syllable durations produced by the implant users.) The average of the three regression slopes for the voiceless plosives was 0.19, and 0.03 for the voiced. These results are reasonably consistent with determinations based on several other studies.¹

In the next step, pre- and postactivation tokens (i.e., those from the first two and the last two recordings, respectively) were paired off by voicing, place of articulation, and order in the protocol. The change in syllable duration from pre- to postactivation was computed, multiplied by the voiceless or voiced regression slope, and added to the postactivation VOT to obtain VOT_c , that is, postactivation VOT corrected for rate. Each pair of tokens, pre- to post-, had associated with it, then, a change in VOT (the postactivation VOT_c minus the preactivation VOT) and a change in each of the three postural correlates, SPL, $F0$, and $H1-H2$.

B. Experiment II: Short-term modifications

1. Subjects

Two of the subjects from experiment I (FA and FB) participated in experiment II. This study was conducted approximately 2 years postactivation of the implant speech processor for subject FA, and 1 year postactivation for subject FB. Results from a longitudinal study of vowel production (cf. Perkell *et al.*, 1992) in which these speakers served indicate that many, but not all, speech parameters had reached stable values for these subjects (referred to as $F1$ and $F2$ in that study) by the time recordings were made for the present study. Implant users were asked to stop wearing their speech processors 24 h before coming to the laboratory for experiment II. This intervention made it impossible for speakers to hear their own speech, that of others, and the transmission conditions, during those waking hours in which they spoke. The interruption of hearing was several orders of magnitude shorter than their prior experience with deafness but it was considerably longer than the brief interruptions (usually with masking) that have been practiced in laboratory experiments.

2. Speech elicitation

The elicitation set consisted of five blocks, each of which contained ten repetitions of the voiced and ten of the

TABLE I. Effects of activating a cochlear implant speech processor on perception of voicing and production of plosive VOT and concurrent acoustic correlates of speech postures. The table reports percent errors in assigning voicing in a consonant identification test, years postactivation of the speaker's cochlear implant speech processor when the test was administered, mean VOTc before and after processor activation, statistically significant increases in VOTc and significant changes in three correlates of speech posture (*t* test, $p < 0.05$). Means for five speakers in a longitudinal study and for two of those speakers in a short-term stimulus modification study. Each longitudinal mean is based on typically 18 tokens (30 tokens for MD. $H1-H2$ was measurable on six voiced tokens of FC and seven of MC). Each short-term mean is based on typically ten tokens in the first processor-off block (PRE) and in the consecutive processor-on block (POST). (For the short-term study, PRE refers to the initial processor-off condition and POST to the two consecutive processor-on conditions pooled.)

Implant user	Longitudinal paradigm					Short-term		
	FA	FB	FC	MD	MC	FA	FB	
% voicing errors	10	3	5	6	4	34	8	
Years postactivation	6	5	3	2	4	2	1	
Voiceless								
VOTc PRE	ms	54	103	95	71	73	82	101
POST	ms	71	112	103	81	69	93	82
POST-PRE	ms	17	8	7	9			
POST-PRE SPL	dB	-8	-3	-2	-6		-3	-7
F_0	Hz	-104		-7	-13	10	-25	
$H1-H2$	dB		5	-1	4			2
Voiced								
VOTc PRE	ms	12	12	17	20	13	12	12
POST	ms	25	17	16	23	11	18	15
POST-PRE	ms	13	5		3		6	3
POST-PRE SPL	dB	-11	-3		-4	3	-6	-4
F_0	Hz	-78		-9	-10		-44	
$H1-H2$	dB		3		3			3

voiceless alveolar plosive within the carrier phrase "It's a /Cad/ again." These two utterances were randomized with other materials not analyzed in the present study; vowel productions by FA, FB, and another subject were analyzed by Svirsky *et al.* (1992). The five blocks were read sequentially under the following conditions: (1) with the speech processor remaining off, after the 24-h deprivation period, (2) immediately after turning the speech processor on, (3) with the speech processor still on, approximately 15 min after it was turned on, (4) immediately after turning the speech processor off again, and (5) with the speech processor still off, approximately 15 min after it was turned off. Thus speech processor status in these five blocks was (1) OFF, (2) ON, (3) ON, (4) OFF, and (5) OFF.

3. Recording, calibration, signal processing, and data analysis

Recording, calibration, signal processing, and data extraction were the same as for experiment I.

4. Correction for speaking rate

The VOT measurements for the tokens recorded in each of the two ON conditions of experiment II were corrected for rate as follows. The tokens from each ON condition were paired off with the set in the first OFF condition used as a baseline. The change in syllable duration in each pair from baseline to ON token was computed, multiplied by the voiceless or voiced regression slope, and added to the ON token VOT to obtain VOTc, that is, processor-on VOT corrected for rate. Each pair of tokens, baseline to ON, had associated

with it, then, a change in VOT (the ON VOTc minus the baseline VOT) and a change in each of the three postural correlates, SPL, F_0 , and $H1-H2$.

II. RESULTS AND DISCUSSION

A. Effects of processor activation on VOT and postural correlates

Table I shows values of VOTc pre- and postactivation of the subjects' implant speech processors. In the longitudinal study (first five columns), all speakers except MC significantly increased voiceless or voiced VOTc after extended exposure to their own voices and those of others (increases are labeled POST-PRE; only significant differences, determined with *t* tests for matched pairs, $p < 0.05$, are shown).

All of the implant users were able to assign the voicing feature correctly in perceptual tests when using their cochlear implants. The first row of data in Table I presents each speaker's error rate in assigning plosive voicing during the consonant identification test closest in time to the longitudinal or short-term experiment (scores, corrected for guessing, from Lane *et al.*, 1994, except for speaker MD). Significant increases in VOTc ranged from 7 to 17 ms for the voiceless plosives, and 3 to 13 ms for the voiced. FA, FB, and MD increased both voiceless and voiced VOTc significantly; FC increased only voiceless VOTc and MC increased neither. In the short-term study, significant increases in VOTc were found only for the voiced plosives. (For this paradigm, "PRE" refers to the initial processor-off condition and "POST" to the two consecutive processor-on conditions pooled.)

Table I shows that every significant VOTc increase fol-

lowing processor activation was accompanied by one or more significant changes in postural correlates. SPL characteristically dropped following activation; in only one instance was there a significant increase: Speaker MC increased voiced token SPL 3 dB. Several investigators have reported that postlingually deafened adults speak excessively loudly (e.g., Leder *et al.*, 1987b), and Perkell *et al.* (1992) found that the vowel inventory was read more softly by all four of their speakers after they began receiving prosthetic hearing.

F_0 fell significantly with processor activation in vowels following both voiceless and voiced plosives obtained from three speakers (FA, FC, MD) in the longitudinal study and from FA in the short-term study. There was one significant increase in F_0 , observed with MC's voiceless plosives. Speaker FA reduced F_0 pre- to postactivation from 255 to 164 Hz (vowels following voiceless and voiced plosives pooled). (Relative to normative data reported for female voices by Holmberg *et al.*, 1988, F_0 fell from 2.08 standard deviations above the mean to 1.7 below it.) As Table I shows, this drop in F_0 was accompanied by an average drop in vowel SPL for voiced and voiceless syllables pooled of 9.5 dB. Speakers FC and MD showed small but reliable decreases in average F_0 . FB, on the contrary, showed no reliable F_0 change following either voiceless or voiced plosives; her average F_0 preimplant was 183 Hz. Other investigators have reported changes in F_0 with implant use, e.g., Kirk and Edgerton (1983), Plant and Oster (1986), Ball and Faulkner (1989).

With prolonged processor use, two speakers increased H_1-H_2 on vowels following both voiced and voiceless plosives (FB did so in both experimental paradigms; MD served only in the longitudinal study). Speakers FA and MC did not change H_1-H_2 significantly and FC decreased H_1-H_2 on voiceless tokens very slightly, 1 dB.

In accord with the theoretical position that audition has the dual roles of validating phonemic settings and monitoring transmission conditions, the first hypothesis predicted that, with some hearing restored, increases in VOT would be accompanied by changes in three postural correlates. Some deafened speakers did indeed increase VOTc when some hearing was restored with cochlear prosthesis, and they also made many concurrent changes in H_1-H_2 , SPL, and F_0 . However, as we have explained, VOTc changes may be traceable, at least some of the time, to concurrent changes in posture.

B. Discriminability of voicing for the implant user

The short-term stimulus modification experiment, conducted with FA and FB, 2 and 1 years, respectively, after processor activation, occurred at a time when their error rates in voicing assignment on the phoneme identification test were 34% and 8%, respectively. By the time of their two final longitudinal recordings ("POST") 6 and 5 years postactivation, respectively, their error rates had fallen to 10% and 3%. The remaining implant users had error rates of 6% or less.

Inference from the low error rates in voicing assignment shown in Table I suggests that the voicing feature was quite

audible to our implant users. In general, studies of speech perception by users of the Ineraid implant indicate that the voicing feature is transmitted well (see the review in Rabinowitz *et al.*, 1992). With an Ineraid user, Dorman *et al.* (1988) found normal labeling functions on the synthetic VOT continuum and a normal boundary shift as a function of varying first-formant onset.

There need not be a one-to-one correspondence between the acoustic correlates of the voicing feature that our subjects attend to and those that they end up changing as a result of modifying underlying articulation. A change in VOT—that is, the measured interval between the burst and the start of periodicity—is an indicant of a larger set of related changes. The amplitude envelope for voiced stops is characterized by a rapid rise following the release burst, whereas for voiceless stops the release is followed by a long low amplitude aspirated segment (Dorman *et al.*, 1990). In the latter case, formant transitions are usually completed before onset of the following vowel, whereas in the voiced plosives there is rapid spectral change at the onset of voicing (Stevens and Klatt, 1974). When voicing is delayed and there is a "cut-back" of F_1 , the F_1 onset frequency and the duration and frequency range of the F_1 transition are also affected. Fundamental frequency contour, burst intensity, and aspiration are all additional cues to voicing. Any one or more of these covarying cues might be the basis of auditory validation of phonemic settings for voicing, leading the implant user to change VOT (and other correlates of the voicing contrast).

C. Effects of short-term changes in processor state

The changes speakers made to VOTc and to postural correlates with changing processor state in the short-term study were, on the whole, consistent with those found in the longitudinal study, although the short-term experiments for FA and FB were conducted 2 and 1 years, respectively, postactivation of their implant speech processors, whereas their final recording sessions for the longitudinal study were approximately 6 and 5 years postactivation (Table I, second row of data). Figures 1–4 present, for speakers FA and FB, the changes in mean VOTc, SPL, F_0 , and H_1-H_2 , respectively, that occurred when their speech processors were turned off, on, and off again. Also plotted in Fig. 1 for reference are mean VOT in each of the two longitudinal preactivation baseline sessions (marked "PRE-1" and "PRE-2") and mean VOTc in the most recent longitudinal session ("REC"). Mean values of VOTc or of postural parameters of the voiceless plosive (V^-) are shown in the top panels, voiced (V^+) in the bottom; vertical bars are one standard error about the mean. Asterisks indicate a significant difference in the mean parameter value contrasting the initial processor-off block with the two following processor-on blocks, or contrasting the two processor-on blocks with the final two, processor-off blocks (ANOVA with planned comparisons, $p < 0.05$; significant changes from ON1 to ON2 and from OFF2 to OFF3 conditions are not indicated).

Although speaker FA increased voiceless VOTc longitudinally a significant 17 ms (from preactivation baseline to final recording sessions, Table I), her increase of 10 ms in the short-term experiment (Fig. 1, upper left panel, OFF1 to

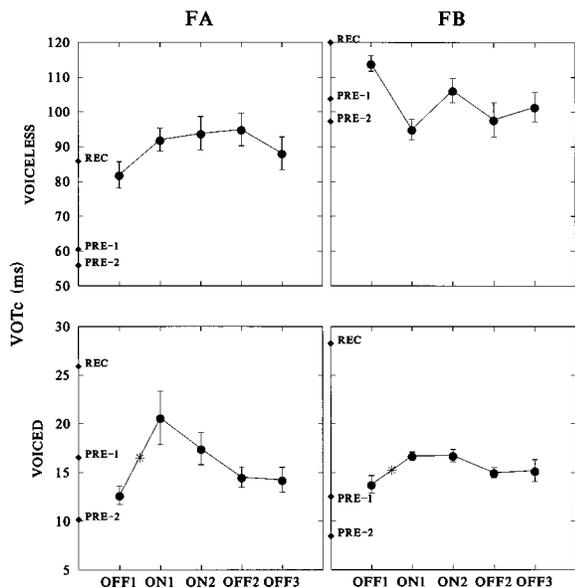


FIG. 1. Mean VOTc (VOT corrected for rate) as a function of processor state in two implant users serving in a short-term stimulus modification experiment. Subjects turned off their processors for 24 h before coming to the laboratory, where they read an elicitation set once with their processors left off, twice with them turned on, and twice with them turned off again. Mean values of VOTc of the voiceless plosive are shown in the top panels, voiced in the bottom; vertical bars are one standard error about the mean. Each point is the average of ten determinations. PRE-1 and PRE-2 indicate average values of VOTc in each of the two preactivation sessions. REC is its mean value in the most recent longitudinal recording before the short-term experiment. Asterisks indicate a significant difference in the mean parameter value contrasting the initial processor-off block with the two following processor-on blocks, or contrasting the two processor-on blocks with the final two, processor-off blocks ($p < 0.05$).

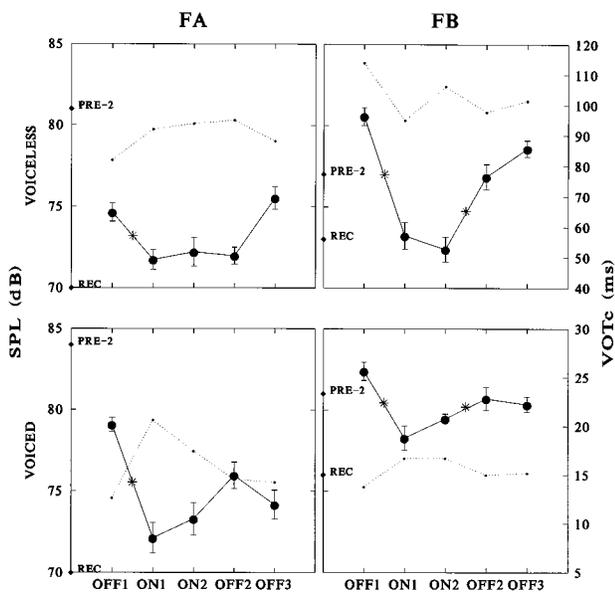


FIG. 2. Mean value of vowel SPL (dB) as a function of processor state in two implant users. Plotting conventions as in Fig. 1. PRE-2 marks the average value of the parameter in the second baseline recording preimplant. To facilitate comparing changes in postural correlates with the direction of changes in VOTc introduced by changing processor state, the functions from Fig. 1 are reproduced. The vertical spacing between the two plotted functions is arbitrary.

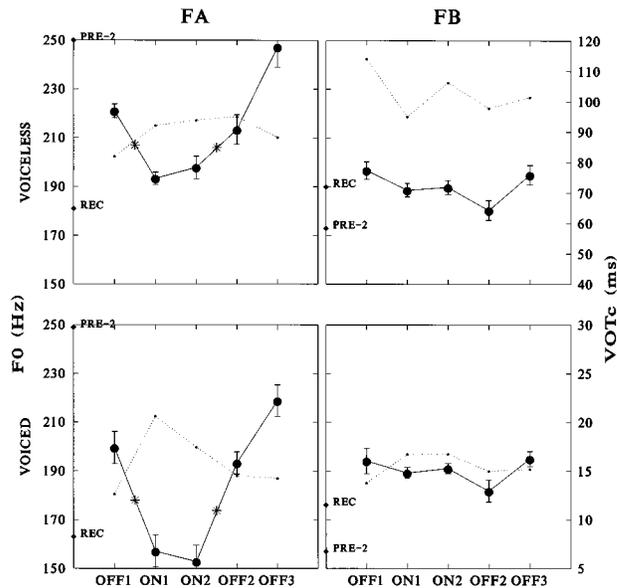


FIG. 3. Mean value of vowel F_0 (Hz) as a function of processor state in two implant users. Plotting conventions as in Fig. 2.

ON1+ON2 pooled), when her processor was turned on after 24 h of nonuse, was not statistically reliable. Likewise, speaker FB increased voiceless VOTc longitudinally but not in the short term: indeed, here her voiceless VOTc fell from OFF1 to ON1+ON2. On the other hand, a reliable increase in *voiced* VOTc (Fig. 1, lower panels) was obtained for each speaker, when the implant speech processor was turned on (Table I; FA: $F[1,7]=7.1$, $p < 0.05$; FB: $F[1,7]=10.0$, $p < 0.05$). It is noteworthy (Fig. 1) that voiceless VOTc, which did not increase significantly when processors were turned on in the short-term study, had not decreased significantly when measured after 24 h with the processor turned off. In contrast, voiced VOTc, which did increase when the

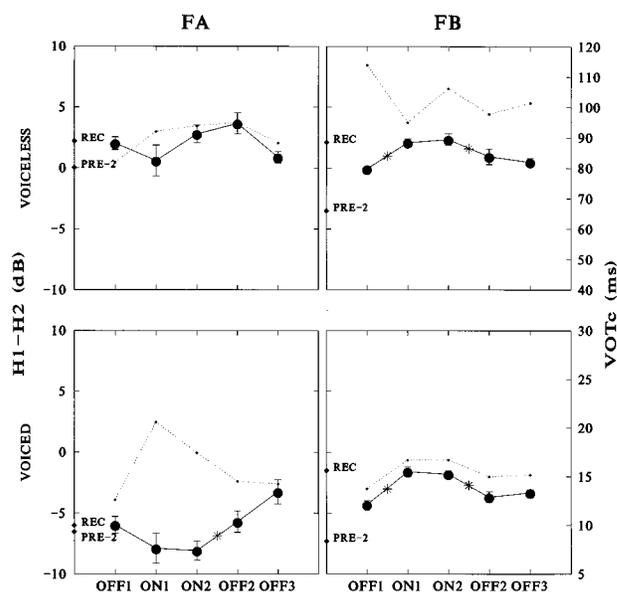


FIG. 4. Mean value of vowel H_1-H_2 (dB), a correlate of breathiness, as a function of processor state in two implant users. Plotting conventions as in Fig. 2.

processor was turned on, had fallen to preimplant baselines when measured after 24 h with the processor off.

The effects on the postural correlates of turning the speaker's speech processor off and on for short periods are presented in Figs. 2–4. To facilitate comparing changes in postural parameters with the direction of changes in VOTc introduced by changing processor state, the VOTc functions from Fig. 1 are reproduced by dotted lines in the subsequent figures, with scales given on the right-hand vertical axes. PRE-2 marks the average value of the postural correlate in the second baseline recording preimplant and REC its value in the most recent longitudinal recording before the short-term experiment.

1. SPL

The changes in SPL produced during the short-term modification study (Fig. 2) were consistent with those found in the longitudinal study. Prolonged implant use longitudinally led FA to lower vowel SPL following voiceless plosives by 8 dB and following voiced plosives by 11 dB (Table I). In the first recording made after the speech processor of her implant had been off for 24 h (Fig. 2, left-hand panels), FA's SPL had climbed substantially above values obtained in the most recent longitudinal recording (labeled REC in the figure) and toward preimplant values (PRE-2). Turning the speech processor on caused FA to lower her SPL (toward REC) significantly ($V-:F[1,7]=6.7, p<0.05$; $V+:F[1,7]=28.8, p<0.01$), consistent with the longitudinal drop in SPL from pre- to postactivation. Turning the processor off again had no immediate significant effect on vowel SPL following voiced or voiceless plosives, but by the end of the experiment, with the processor still off, her SPL rose reliably 3 dB on the voiceless tokens (Tukey HSD, $p<0.01$).

FB, in the longitudinal study, significantly lowered her SPL by 3 dB from preimplant to the final two recording sessions (Table I). The drop from preimplant (PRE-2, right-hand panels, Fig. 2) to the earlier recording session closest to the short-term experiment (REC) was larger, 4.4 dB ($V-:t[8]=4.3, p<0.01$; $V+:t[8]=4.5, p<0.01$). After 24 h with her speech processor off, her SPL rose and exceeded preimplant values (cf. PRE-2, Fig. 2). When the speech processor of her cochlear implant was then turned on again, FB promptly lowered SPL again significantly on both voiceless and voiced tokens, dropping to recent values for the voiceless syllables and toward recent values for the voiced ($V-:F[1,9]=132.1, p<0.01$; $V+:F[1,9]=35.8, p<0.01$). Finally, when the processor was turned off again, SPL rose significantly for both voiceless and voiced tokens ($V-:F[1,9]=64.3, p<0.01$; $V+:F[1,9]=8.4, p<0.05$).

Thus, for FB as for FA, turning the speech processor of her cochlear implant off for 24 h led to increases in the SPL of both voiceless and voiced tokens. Turning it back on after 24 h led to decreases in SPL for both voiceless and voiced tokens, as did processor activation in the longitudinal study. The complementary pattern in the short-term experiment was observed for VOTc in Fig. 1 for voiced but not voiceless tokens.

2. F0

Four of the five speakers changed F0 reliably after extended processor use (Table I). FA reduced F0 an average of 91 Hz. The effects on F0 of turning off the speech processor of her cochlear implant for 24 h are shown in Fig. 3. F0 of vowels following both voiceless and voiced plosives increased from recent values in longitudinal recordings an average of 39 Hz ($V-:F[1,7]=39.8, p<0.01$; $V+:F[1,7]=16.4, p<0.01$). When the processor was turned on again, F0 fell; when it was then turned off, F0 rose ($V-:F[1,7]=42.1, p<0.01$; $V+:F[1,7]=26.6, p<0.01$). On the other hand, FB's F0 did not change significantly between baseline and final sessions (Table I) nor when her processor state was varied (Fig. 3). Thus this postural correlate also varies with processor state in the short-term experiment in a way consistent, for each speaker, with her changes in F0 in the longitudinal study.

3. H1-H2

Recall that, for the vowels following voiceless and voiced plosives of FA, there was no difference in H1-H2 between preimplant and final sessions 4 years later (Table I). Similarly, H1-H2 did not change reliably between preimplant baseline (Fig. 4, left) and the recording closest in time to her short-term experiment (REC). Consistent with this finding, turning the speech processor of her implant off for 24 h had no significant effect on her H1-H2 (the first plotted symbol is at the level of "REC"), nor did turning it on and then off again later in the same recording session, except for a significant increase in H1-H2 for voiced tokens between the two ON and the final two OFF conditions. Thus FA's significant VOTc increase for the voiced plosives (from OFF1 to ON1+ON2 pooled) was not accompanied by a significant change in H1-H2.

Subject FB did increase H1-H2 significantly by an average of 4 dB after more than a year of implant use (Table I). The effects of short-term modifications of her hearing proved to be consistent with the corresponding longitudinal changes. Arriving in the laboratory after 24 h without hearing, FB uttered voiceless and voiced plosives with a mean value of H1-H2 (OFF1) that was 2.3 dB lower than in the recent recording with the processor on (REC), but 3.3 dB higher than the preimplant baseline a year earlier (Fig. 4, right, PRE-2). Turning the processor on led to an increase in H1-H2 ($V-:F[1,9]=93.4, p<0.01$; $V+:F[1,9]=68.1, p<0.01$). Turning the processor off again led to significant reductions in H1-H2 ($V-:F[1,9]=22.8, p<0.01$; $V+:F[1,9]=62.4, p<0.01$). For FB's voiced plosives (but not voiceless), trends in mean VOTc and H1-H2 across the conditions of the short-term experiment appear to correspond; the question of whether they are causally related is discussed below.

Consistent with the role assigned to audition of monitoring transmission conditions, the second hypothesis provided that, when speakers' processors are turned off for brief periods, acoustic correlates of posture will revert toward their preimplant values and when processors are turned on again, postural correlates will recover recent average values. The

TABLE II. VOTc increases significantly correlated with concurrent postural changes in two implant users and one speaker with normal hearing. The table reproduces the change in VOTc from baseline to final sessions (ms) shown in Table I and presents significant contributions to predicted VOTc change from changes in each of three postural correlates (ms) and the value of the additive constant k (ms) when regressing changes in VOTc changes on changes in the acoustic correlates of speech postures.

Subject Paradigm	FA (voiced)	FB (voiced)		Hearing (voiced)
	longitudinal	longitudinal	short term	instructed
R	0.65	0.66	0.68	0.39
POST-PRE VOTc (ms)	13	5	3	2
ms contribution from:				
SPL	14	2	4	-5
$F0$				
$H1-H2$		-3		
k		7		6

prediction is confirmed for SPL for both speakers; for $F0$, for the one speaker who showed significant $F0$ change longitudinally; and for $H1-H2$, for the one speaker who showed significant $H1-H2$ change longitudinally.

D. Postural mediation of VOTc increases

In order to examine the contributions of postural changes to the observed VOTc increases whose averages are shown in Table I, multiple regressions were performed with measures of postural changes as the independent variables and VOTc increases as the dependent variable (see Table II). Tokens analyzed were from the initial two and the final two longitudinal recording sessions, respectively, or from the first processor-off condition and the consecutive two processor-on conditions in the short-term study. Tokens uttered with the implant speech processor off were paired off with tokens uttered with the processor on by voicing, place of articulation, and order in the elicitation set, and the differences in VOTc, $H1-H2$, SPL, and $F0$ were computed. (When computing changes in the short-term study, the first OFF condition served as a baseline for tokens in both of the two ON conditions.) Finally, the VOTc differences were regressed on their associated changes in postural correlates, separately for each experiment (longitudinal and short term), subject (FA, FB), and type of syllable (voiceless and voiced).

Increases in voiceless plosive VOTc were never associated with concurrent changes in postural correlates in the short-term experiment and for only one of the five speakers in the longitudinal experiment, MD. A 13-Hz reduction in MD's $F0$ contributed 17 ms to the predicted increase in VOTc ($R=0.56$, $F[3,24]=3.6$, $p<0.05$). This inverse linkage between VOTc and $F0$ is not consistent with the earlier discussion of laryngeal mechanics; it may be mediated by the covariation of $F0$ with SPL; the latter fell 6 dB from pre- to postactivation for this speaker. The partial correlation of $F0$ and SPL changes with $H1-H2$ held constant in MD's voiceless tokens is ($r[17]=0.73$, $p<0.01$).

Turning to the voiced tokens, when VOTc changes were regressed on concurrent changes in postural correlates, three significant multiple correlations (R) were obtained, two in the longitudinal study, for FA and FB, and one in the short-term study (FB). (Findings in the right-hand column obtained from a normally hearing speaker instructed to make postural

changes are discussed below.) Table II shows the significant increases in voiced VOTc from baseline to final sessions (row 2 of data) and the significant multiple correlations obtained (row 1). Rows 3–6 report significant additive components of the regression equations; they present the contributions in ms to predicted VOTc arising from changes in each of three postural correlates, and the additive constant k .

It is evident that SPL decreases in longitudinal and short-term experiments, reported in Table I, contributed significantly to predicted voiced VOTc increases for each one of the significant multiple R 's shown in Table II. For FA's voiced plosives (Table II, col. 1), an 11-dB drop in SPL (Table I) contributed 14 ms, approximately the 13-ms increase in predicted VOTc. For speaker FB (Table II, cols. 2 and 3), a 3-dB drop in SPL contributed 2 ms to predicted voiced VOTc increases in the longitudinal study, and a 4-dB drop in SPL contributed a significant 4 ms to predicting the observed VOTc increase of 3 ms in the short-term study. The inverse relation between SPL changes and predicted VOTc changes found with both speakers is consistent with our earlier discussion of laryngeal mechanics, to which we return below.

Finally, $H1-H2$ made a significant contribution to predicted VOTc increases in only one case, voiced tokens from FB. (However, $H1-H2$ was not measurable on 12 of 18 voiced tokens of FC and 11 of 18 for MC, and this may have contributed to this negative outcome.) Speaker FB increased voiced plosive VOTc 5 ms (Table II, col. 2, row 2). Looking down the column, the largest component of the predicted increase came from the additive constant k of 7 ms, and the reduction in SPL contributed 2 ms. Finally, the 3-dB increase in speaker FB's voiced plosive $H1-H2$ reduced predicted VOTc by 3 ms. This is the only set of tokens in the experiment in which a relation was found ($r[17]=-0.57$, $p<0.05$) between changes in $H1-H2$ and those in VOTc. Once again, the drop in SPL may be mediating the correlation between this postural correlate and VOTc. Indeed, an increased glottal aperture (associated with an increase in $H1-H2$) may be one mechanism for implementing a decrease in SPL.

1. Instructed changes in posture and relation to VOTc

In order to gauge in a preliminary way the size of the changes in VOTc that might be expected from changes in posture, we asked one female speaker with normal hearing to

read the same elicitation phrases as in the short-term modification study while adjusting (perceived) SPL, F_0 , and breathiness to one of three levels—normal, intermediate, or high. The speaker was instructed to attempt to hold each postural correlate constant in turn while varying the other two over intermediate and then high levels. This yielded nine experimental conditions with each of the three variables held constant in turn, for a total of 27 conditions. There were 10 trials in each condition, for a total of 270 trials with the voiceless alveolar plosive and the same number with the voiced plosive. Parameter values measured when the speaker aimed to produce syllables with normal $H1-H2$, SPL, and F_0 (30 trials) were used as a baseline. The comparison values came from the subset of trials when the speaker aimed to hold two parameters at normal levels while varying the other one over intermediate and high levels (60 trials). Values of SPL, F_0 , $H1-H2$, duration, and VOT measured on the baseline tokens were paired off with and then subtracted from those of the comparison tokens, first at intermediate level, then at high, and the VOTc increment for each pair was computed. Finally, the VOTc increments from baseline to comparison were regressed on the associated changes in the acoustic correlates of posture.

Although the speaker with normal hearing varied $H1-H2$ over 5 dB, SPL over 8 dB, and F_0 over 42 Hz from normal to high levels, those changes in correlates of posture were associated with only a nonsignificant 3.5-ms drop in VOTc among the voiceless tokens. There was a small but significant increase in VOTc among the voiced tokens, as shown in Table II (last column). The sole postural correlate making a significant contribution to predicting that voiced VOTc increase was SPL; the relation was inverse, as was also observed among the implant users. The scale factor (regression coefficient) was -0.4 , indicating a 2.4-ms contribution to predicted VOTc increase per 6-dB decrease in SPL. Scale factors for the other postural correlates were not reliably different from zero.

2. Summary

The third hypothesis provided that the postural changes accompanying VOTc changes with the restoration of some hearing will frequently be the source of the VOTc changes as a result of linkages between speech postures and the phonatory gestures determining VOT. This hypothesis finds some support from the four significant multiple regressions in the present study. One such regression was obtained for the voiceless tokens of speaker MD; the other three were for voiced tokens measured in the longitudinal (FA, FB) and short-term (FB) experiments. VOT among voiced plosives may be more vulnerable to influence by speech postures than among voiceless plosives since, as discussed earlier, VOT in the voiced plosives may depend primarily on the timing of the drop in supraglottal overpressure, which is determined in part by respiratory and laryngeal postures. Voiceless plosives, on the other hand, are characterized by an active abduction of the vocal folds.

Although there were, then, four significant multiple regressions of VOTc increases on changes in postural correlates, five other significant increases in VOTc (voiceless to-

kens for FA, FB, and FC longitudinally and voiced for MD in the longitudinal study and FA in the short-term study) were not correlated with postural changes. The absence of significant correlations between increases in VOTc and changes in postural correlates in those cases does not, of course, confirm that those increases were solely the result of phonemic resettings due to auditory validation. However, the fact that such correlations were found for some speakers under the conditions of this experiment is consistent with the theory that provides two distinct mechanisms for increases in VOTc with prosthetic hearing: linkage to postural adjustments due to transmission monitoring and direct phonemic resetting due to auditory validation. Two findings in this study lead us to suggest that VOTc changes in voiced plosives with changes in hearing status may reflect predominantly postural changes while VOTc changes in voiceless plosives may reflect predominantly auditory validation of phonemic settings. The changes in voiced VOTc yielded high multiple correlations with concurrent postural changes, especially SPL. And changes in voiceless VOTc were observed over a period of years in the longitudinal study but not with the shorter changes in hearing status in the short-term study.

This conclusion must remain very tentative without convergent evidence from other experimental methods. We have begun to examine measures of implant users' ability to discriminate between synthesized plosive-vowel syllables containing their own average preimplant values of VOT and those of speakers with normal hearing. Where implant users are unable to make those discriminations and yet increase VOT with processor use, a postural contribution to VOT may be implicated. One limitation of this approach is that auditory validation of the voicing contrast may not be controlled directly by perceived VOT, as explained above, so that performance on the synthetic speech discrimination task will be predictive only to the extent that VOT is correlated there with other cues controlling validation. Another approach to gauging the contribution of postural changes to VOT changes would be to extend the preliminary normative study reported here not only to more normally hearing speakers instructed to vary the several parameters but also to implant users similarly instructed. A regression model of linkages obtained in this way could assist in separating postural changes from phonemic resetting. If the VOT increases obtained from implant users postactivation were considerably larger than the increases predicted when their postural changes were entered in their regression model of linkages, auditory validation of phonemic synergisms would be implicated.

ACKNOWLEDGMENTS

We are grateful to Dr. William Rabinowitz, Research Laboratory of Electronics, Massachusetts Institute of Technology, for providing us with the results of the consonant identification tests and examining them with us. Dr. Joanne Miller, Department of Psychology, Northeastern University, kindly gave us the data from normally hearing speakers' VOTs for analysis. We are indebted to the Harvard University Psychology department "brown bag" statistical lunch for discussions of data analysis, in particular to Dr. Robert

Rosenthal, Dr. Donald Rubin, and Dr. Douwe Yntema. We appreciate the helpful criticisms of an earlier version of the article given by Dr. Arlene Carney, University of Minnesota, and Dr. Anders Löfqvist, Haskins Laboratories, and by two anonymous reviewers. This work was supported by N.I.D.C.D. Grant No. DC 00361 to the Massachusetts Eye and Ear Infirmary, the Massachusetts Institute of Technology, and Northeastern University, Dr. Joseph B. Nadol and Dr. Donald Eddington, Co-Principal Investigators.

¹Volaitis and Miller (1992) had three speakers produce runs of the six English plosives in C/i/ context at eight different rates of speech; runs were repeated four times. We computed the slopes of the lines of best fit relating VOT to syllable duration for each place of articulation, value of voicing, and speaker, using the subset of syllables with durations between 200 and 400 ms. The mean slope for voiceless plosives was 0.19 and for voiced, 0.03; their standard deviations were 0.12 and 0.02. Since all implant users in the present study shortened mean syllable duration with processor activation, steeper correction slopes would add larger corrections to their post-activation VOTs, and more shallow slopes would add smaller corrections.

Summerfield (1975) had six male speakers give 15 repetitions at three rates of speech of "Why are you a C1V1 when you're a C2V2," where C=/p/ or /k/. Summerfield reported averages across speakers of VOT and the durations of their preceding carrier phrases. [Although he found that place of articulation of the initial stop interacted with the following vowel in determining VOT, Port and Rotunno (1979) did not replicate this finding.] We estimated mean syllable durations by dividing the reported durations of the carrier phrases by the number of syllables in the phrases. The slopes of the lines of best fit relating mean VOT to mean syllable duration in the preceding carrier were computed for each of two places of articulation and two positions in the carrier phrase. The mean slope for the voiceless plosives was 0.22, the standard deviation 0.083.

Pind (1995) had four Icelandic speakers read word lists at five rates. The slope of the line of best fit relating VOT to syllable duration for the voiceless aspirated alveopalatal plosive was 0.17 over the range 250–500 ms. There was no significant slope for the voiceless nonaspirated alveopalatal plosive.

Port and Rotunno (1979) studied the effect of the final consonant (experiment I) and of vowel intensity (experiment II) on the VOT of syllable-initial voiceless plosives in CVC(C) syllables. Average measures of VOT and vowel duration are reported. In experiment I, eight speakers read six syllables (3 plosives×2 endings) five times each. In experiment II, five speakers read 18 syllables (3 plosives×3 tense/lax vowel pairs) five times each. We added average VOT and vowel duration measures to obtain mean syllable duration, and computed slopes relating VOT to syllable duration. The effect of replacing a syllable-final consonant cluster (/pt/) with a nasal consonant (/n/) in experiment I was to reduce both VOT and syllable duration; the three slopes for the three places of articulation were 0.20, 0.14, and 0.18 (/p,t,k/). The effect of tensing the vowel in experiment II was to increase both VOT and syllable duration; the mean of the nine slopes was 0.13, the standard deviation 0.047. Diehl *et al.* (1980) had eight speakers read a carrier phrase with the test word /ka/ five times each at fast and slow rates. The change in VOT that correspond to the change in syllable duration was slope =0.2 for the males and 0.29 for the females.

Ball, V., and Faulkner, A. (1989). "Speech production of postlingually deafened adults using electrical and acoustic speech pattern prostheses," *Speech, Hearing and Language: Work in Progress*, Report of the University of College of London, Vol. 3, pp. 13–32.

Diehl, R. L., Souther, A. F., and Convis, C. L. (1980). "Conditions on rate normalization in speech perception," *Percept. Psychophys.* **27**, 435–443.

Dorman, M., Hanley, M., McCandless, G., and Smith, L. (1988). "Auditory/phonetic categorization with the Symbion multichannel cochlear implant," *J. Acoust. Soc. Am.* **84**, 501–510.

Dorman, M. F., Soli, S., Dankowski, K., Smith, L. M., McCandless, G., and Parkin, J. (1990). "Acoustic cues for consonant identification by patients who use the Ineraid cochlear implant," *J. Acoust. Soc. Am.* **88**, 2074–2079.

Hirose, H. (1976). "Posterior cricoarytenoid as a speech muscle," *Ann. Otol. Rhinol. Laryngol.* **85**, 334–343.

Holmberg, E., Hillman, R. E., and Perkell, J. S. (1988). "Glottal airflow and

transglottal air pressure measurements for male and female speakers in soft, normal and loud voice," *J. Acoust. Soc. Am.* **84**, 511–529.

Holmberg, E., Hillman, R., Perkell, J. S., and Gress, C. (1994). "Relationships between intra-speaker variation in aerodynamic measures of voice production and variation in SPL across repeated recordings," *J. Speech Hear. Res.* **37**, 484–495.

Kirk, K. I., and Edgerton, B. J. (1983). "Effects of cochlear implant use on voice parameters," *Otolaryngol. Clin. North Am.* **16**, 281–292.

Klatt, D. H., and Klatt, L. C. (1990). "Analysis, synthesis and perception of voice quality variations among female and male talkers," *J. Acoust. Soc. Am.* **87**, 820–857.

Lane, H., Wozniak, J., and Perkell, J. (1994). "Changes in voice-onset time in speakers with cochlear implants," *J. Acoust. Soc. Am.* **96**, 56–64.

Leder, S. B., Spitzer, J. B., Kirchner, J. C., Flevaris-Phillips, C., Milner, P., and Richardson, F. (1987a). "Speaking rate of adventitiously deaf male cochlear implant candidates," *J. Acoust. Soc. Am.* **82**, 843–846.

Leder, S. B., Spitzer, J. B., Milner, P., Flevaris-Phillips, C., Kirchner, J. C., and Richardson, F. (1987b). "Voice intensity of prospective cochlear-implant candidates and normal-hearing adult males," *Laryngoscope* **97**, 224–227.

Löfqvist, A. (1980). "Interarticulator programming in stop production," *J. Phon.* **8**, 475–490.

Perkell, J., Holmberg, E., and Hillman, R. (1991). "A system for signal processing and data extraction from aerodynamic, acoustic and electroglottographic signals in the study of voice production," *J. Acoust. Soc. Am.* **89**, 1777–1781.

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**, 2961–2978.

Perkell, J. S., Matthies, M., Svirsky, M., and Jordan, M. (1995). "Goal-based speech motor control: A theoretical framework and some preliminary data," *J. Phon.* **23**, 23–35.

Pind, J. (1995). "Speaking rate, VOT and quantity: The search for higher-order invariants for two Icelandic speech cues," *Percept. Psychophys.* **57**, 291–304.

Plant, G., and Oster, A. (1986). "The effects of cochlear implantation on speech production," *Speech Transmission Laboratory, Quarterly Progress and Status Report (Stockholm)*, Issue 1, pp. 65–84.

Port, R. F., and Rotunno, R. (1979). "Relation between voice-onset time and vowel duration," *J. Acoust. Soc. Am.* **66**, 654–662.

Rabinowitz, W. M., Eddington, D. K., Delhorne, L. A., and Cuneo, P. A. (1992). "Relations among different measures of speech reception in subjects with using a cochlear implant," *J. Acoust. Soc. Am.* **92**, 1869–1881.

Sernicales, W., d'Alimonte, G., and Alegria, J. (1984). "Production and perception on French stops by moderately deaf subjects," *Speech Commun.* **3**, 185–198.

Shannon, R. V. (1993). "Psychophysics," in *Cochlear Implants: Audiologic Foundations*, edited by R. S. Tyler (Singular, San Diego, CA), pp. 357–368.

Stevens, K. N. (1977). "Physics of laryngeal behavior and larynx modes," *Phonetica* **34**, 264–279.

Stevens, K. N., and Klatt, D. (1974). "Role of formant transitions in the voiced–voiceless distinction for stops," *J. Acoust. Soc. Am.* **55**, 653–659.

Summerfield, Q. (1975). "How a full account of segmental perception depends on prosody and vice versa," in *Structure and Process in Speech Perception*, edited by A. Cohen and S. G. Neeboom (Springer–Verlag, New York), pp. 51–68.

Svirsky, M. A., Lane, H., Perkell, J. S., and Wozniak, J. (1992). "Effect of short term auditory deprivation on speech production in adult cochlear implant users," *J. Acoust. Soc. Am.* **92**, 1284–1300.

Tartter, V., Chute, P., and Hellman, S. (1989). "The speech of a postlingually deafened teenager during the first year of use of a multichannel cochlear implant," *J. Acoust. Soc. Am.* **86**, 2113–2121.

Volaitis, L. E., and Miller, J. L. (1992). "Phonetic prototypes: Influence of place of articulation and speaking rate on the internal structure of voicing categories," *J. Acoust. Soc. Am.* **92**, 723–735.

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

Weismer, G. (1984). "Control of the voicing distinction for intervocalic stops and fricatives. Some data and theoretical considerations," *J. Phon.* **8**, 427–438.