

July 01, 1994

Changes in voice-onset time in speakers with cochlear implants

Harlan Lane

Northeastern University; Massachusetts Institute of Technology

Jane Wozniak

Massachusetts Institute of Technology

Joseph Perkell

Massachusetts Institute of Technology

Recommended Citation

Lane, Harlan; Wozniak, Jane; and Perkell, Joseph, "Changes in voice-onset time in speakers with cochlear implants" (1994). *Psychology Faculty Publications*. Paper 5. <http://hdl.handle.net/2047/d20000874>

This work is available open access, hosted by Northeastern University.

Changes in voice-onset time in speakers with cochlear implants

Harlan Lane,^{a),b)} Jane Wozniak,^{a)} and Joseph Perkell

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

(Received 24 May 1993; accepted for publication 25 March 1994)

Voice-onset time (VOT) and syllable duration were measured for the English plosives in /Cad/ (C=consonant) context spoken by four postlingually deafened recipients of multichannel (Ineraid) cochlear implants. Recordings were made of their speech before, and at intervals following, activation of the speech processors of their implants. Three patients reduced mean syllable duration following activation. Using measures of VOT and syllable duration from speakers with normal hearing [Volaitis and Miller, *J. Acoust. Soc. Am.* **92**, 723–735 (1992)] and from the subjects of this study, VOT is shown to vary approximately linearly with syllable duration over the ranges produced here. Therefore, the VOT of each token was adjusted for the change in syllable duration of that token relative to the mean syllable duration in the first baseline session. This variable, labeled VOTc, was used to evaluate the effects on voicing of the speakers' renewed access to the voicing contrast provided by their implants. Preimplant, all four speakers characteristically uttered voiced plosives with too-short VOT, compared to the measures for hearing speakers. Voiceless plosive mean VOT was also abnormally short for two of the speakers, and close to normal for the remaining two. With some hearing restored, subjects made relatively few errors with respect to voicing when identifying plosives in listening tests, and three of the four speakers lengthened VOTc. The findings are interpreted as supporting the hypothesis that speakers use their hearing to calibrate mechanisms of speech production by monitoring the relations between their articulations and their acoustic output.

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

INTRODUCTION

In order to clarify the role of hearing in speech production, we have been examining the physiological and acoustic properties of the speech of deafened adults before and after they recover some hearing with a cochlear implant (Lane *et al.*, 1991; Perkell *et al.*, 1992; Svirsky *et al.*, 1992). When speakers who became profoundly, bilaterally, and sensorineurally deaf as adults present anomalous speech patterns, these can likely be attributed to the interruption of their ability to hear themselves and others. If their hearing is then partially restored and their anomalous speech patterns shift toward normal, the implication is that hearing has a role in regulating those speech parameters.

Voice onset time (VOT) must rank high among the segmental properties of speech that invite this kind of investigation of speech regulation by hearing. First, the deaf speaker cannot hear or see the laryngeal gestures that control VOT, and some of the forms of somatosensory feedback that arise from supraglottal articulation may be absent or reduced at the level of the glottis (Davis *et al.*, 1992). Thus if VOT is normally regulated by hearing, it may be particularly vulnerable to deafening; and if it can be discriminated well with the aid of a cochlear implant, it may change following activation of the implant processor. Second, the control of VOT requires precise timing of glottal and supraglottal events and the temporal coordination of articulatory events may be particularly vulnerable to deafening. When Lane *et al.* (1991)

found anomalies in speech breathing in deafened speakers, they suggested that the relative timing of glottal maneuvers might be implicated, among other mechanisms.

Several studies have shown that speakers born deaf or deafened before learning English “blur” the distinction between voiced and voiceless consonants; they fail to implement this phonemic contrast with sufficiently differentiated VOT (Monsen, 1976), airflow (Whitehead and Barefoot, 1980), intraoral air pressure (Hutchinson and Smith, 1976) and laryngeal gestures (Mahshie and Conture, 1983). Levitt and Stromberg (1983), analyzing corpora obtained from congenitally deaf children by Smith (1975) and Gold (1978, 1980), list under “asynchrony errors” numerous substitutions of voiced for voiceless cognates. Taken together, these findings are consistent with the hypothesis that hearing plays a role in regulating VOT, but they do not lend it strong support since early-deafened speakers may not have mastered the phonemic contrast and its implementation during language acquisition. However, Cowie and Douglas-Cowie (1983) transcribed the speech of a group of postlingually deafened adults and report that the opposition between voiced and voiceless consonants was “neutralized” in this population as well. Boothroyd *et al.* (1988) found that five speakers in a group of six postlingually deafened adults were particularly poor at producing the voicing contrast before receiving the Nucleus-22 implant and three of these five improved, according to measures of their intelligibility, following activation of their speech processors. Among the eight segmental contrasts that were studied, speakers using their implants perceived the voicing contrast most accurately, and there was a strong relation between the contrasts they dis-

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

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

criminated accurately and those that improved in production. On the other hand, Tartter *et al.* (1989) found that one postlingually deaf teenager did not reliably change VOT in the plosives after one year of using the Nucleus-22 implant.

The present study examines the voicing contrast in the speech of postlingually deafened implant users in readings made twice before and periodically after activation of their implant processors; in particular, VOT is measured in syllable-initial English plosives varying in manner (voiced/voiceless) and place (bilabial, alveolar, velar). Comparisons of mean VOT are made within subjects, contrasting VOT during prolonged deafness with VOT following the restoration of some hearing, and between our subjects and speakers with normal hearing. This study asks, first: Do late-deafened speakers with prolonged profound deafness present anomalies in VOT of voiced and voiceless plosives? A finding of peculiarly short VOTs among the voiceless plosives or peculiarly long VOTs among the voiced would be consistent with Cowie and Douglas-Cowie's report of frequent phonemic substitutions of voiced and voiceless plosives in this population of speakers, and it would support indirectly the hypothesis that hearing normally plays a role in maintaining the voicing opposition. The study asks, next: Does VOT change systematically with the activation of the cochlear implant processor and, if so, does it change in the direction of values obtained from speakers with normal hearing? Further: Is there evidence that speakers who change values of VOT toward normal after processor activation also discriminate the voicing feature using their implants? If the VOT change is toward values obtained from hearing speakers, and if the implant users discriminate voicing reliably, there is further support for the hypothesis that one of the roles of hearing in normal speech production is to regulate the voicing contrast.

Although such findings would support the hypothesis of a role for hearing in regulating this speech parameter, there are potentially confounding factors that can infirm this support and must therefore be evaluated. Activation of the cochlear implant processor is commonly associated with numerous changes in speech, among them global changes in rate, sound pressure, and fundamental frequency (which we refer to as "postural changes;" see Perkell *et al.*, 1992). When such changes accompany changes in VOT and its discrimination, it is possible that they are responsible for the change in VOT—either solely, or in conjunction with the speaker's newfound ability to discriminate voicing.

I. METHODS

Adult users of cochlear implants are a heterogeneous group with respect to the parameter values of speech. Like hearing speakers of English, they differ in gender, in dialect, and in personal speaking style. In addition, they differ in age at first substantial hearing loss, use of hearing aids, age at deafening, duration of prolonged profound deafness, and adaptation of their speech to these events. Moreover, each implant user had his/her own particular pattern of cochlear damage, implant electrode depth, processor tuning, and ability to interpret stimuli from the processor. Consequently this study employs a single-subject longitudinal experimental design, replicated with four different subjects. Findings are dis-

TABLE I. Subject characteristics.

| Subject | FA | FB | FC | MC |
|---|-----|-----|-----|-----|
| Sex | F | F | F | M |
| Age at implant | 51 | 50 | 47 | 56 |
| Age at onset of profound deafness | 33 | 21 | 47 | 41 |
| Pure tone average loss best ear (dB) | 107 | 104 | 103 | 107 |
| Last score on NU-6 test (% correct) | 8 | 30 | 36 | 6 |
| weeks postactivation | 208 | 108 | 88 | 147 |

cussed separately and statistical tests conducted separately for each subject. Although the implant users in this study differ in many of the ways mentioned, it turns out that they respond similarly in some respects to prolonged deafness and the restoration of some hearing.

A. Implant characteristics, subjects

The Ineraid cochlear implant (Richards Medical Co.) consists of an implanted electrode array, a percutaneous pedestal and connector, and an external sound processor (for a detailed description see: Eddington, 1983; Youngblood and Robinson, 1988). The sound processor has overlapping band-pass filters with crossover frequencies of approximately 0.7, 1.4, and 2.3 kHz. The electrodes, initially spaced 4 mm apart, were successfully positioned 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.

Table I presents information concerning the subjects.¹ Three of the subjects were females. 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. The male subject, MC, had a progressive bilateral hearing loss beginning at age 10 and wore a hearing aid from that time until approximately six months after the activation of the speech processor of his cochlear implant. Consistent with reports that link age at hearing loss to intelligibility in adulthood (Cowie and Douglas-Cowie, 1983), we noted informally that the speech of our subjects was quite intelligible; a formal evaluation has not been conducted, however.

B. Auditory tests

The four subjects had pure tone average losses greater than 102 dB in each ear prior to implant. Two tests from the MAC Battery (Owens *et al.*, 1985), presented aurally prior to implant surgery, were used to assess unaided word recognition. Stimuli were delivered at high speech levels through a speech audiometer with audiometric headphones. The four-choice Spondee test of word recognition was administered to FA unaided, and to FB and MC aided; they scored 35%, 35%, and 40% correct, respectively. FC was tested unaided using the NU-6 open set test and scored 0% and 2% correct

on two administrations. Following activation of their implant processors, all subjects performed above chance on the NU-6 test; their final scores and the number of weeks that had elapsed between the activation of the implant processor and the test administration are shown in the last two rows of Table I. All four patients have continued to use their prostheses regularly.

An index of the speakers' abilities to discriminate plosive voicing using their implants was available from a test of consonant identification administered (for other purposes) at intervals following processor activation (for details, see Rabinowitz *et al.*, 1992). Briefly, each of twelve syllables (six plosives, four fricatives, and two nasals, followed by the vowel /a/) was uttered three times by one male speaker, digitized at a 10-kHz sampling rate, and presented over a loud-speaker in blocks of 72 trials in quasirandom sequences. There were between 18 and 36 presentations of each plosive.

C. Speech elicitation

Two baseline recordings of speech production were obtained from each subject before activating the speech processor of the subject's cochlear implant; the recordings were separated by intervals of 10, 1, and 1 weeks, and 1 day, respectively, for FA, FB, FC, and MC. Some preactivation recordings were made before the implant surgery and others following it. Post-activation 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 subjects did not receive explicit auditory training or speech therapy during the course of this research, with one exception: FC began speech therapy one month prior to the final recording session at 104 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; other speech material was recited for approximately 20 min between each reading.

D. Recording, calibration, signal processing, data analysis

The subject was seated in a comfortable chair 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. The microphone signal was amplified, recorded and later low-pass filtered at 5 kHz and digitized at 10 kHz. Digitization, signal processing, and interactive data extraction were performed with procedures written in the MITSYN languages (Henke, 1989; Perkell *et al.*, 1991) running on a Digital Equipment Corporation engineering workstation. Working with a display of the digitized speech signal of each utterance, an experimenter placed markers at the onset of the plosive release burst (t_1), at the first zero crossing before the onset of periodicity in the

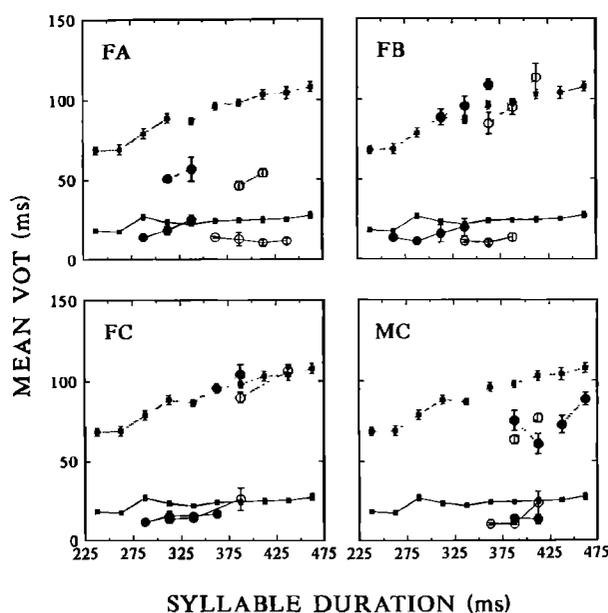


FIG. 1. The relation of VOT to syllable duration. Each panel shows the relation for one speaker between the VOTs produced and the durations of the syllables in which they occurred. Measures were obtained from three hearing speakers (Volaitis and Miller, 1992) (squares) and from postlingually deafened adults before (open circles) and after (filled circles) activation of their implant processors. Mean VOT for voiced (solid lines) and voiceless (dotted lines) plosives are plotted separately as a function of syllable duration in 25-ms class intervals. Vertical bars represent one standard error of the mean (in some cases, the bars are obscured by the symbol for the mean). Each mean for the hearing speakers is based on between 6 and 98 observations; each mean for the deafened speakers between 3 and 13 observations.

waveform (t_2), and at the last zero crossing of the last periodic pulse of the vowel (t_3). VOT was computed as the interval between the first two time markers, and syllable duration as the interval between the first and the third time markers.

A correction was added to each token's VOT to offset the effect of changes in syllable duration (see Sec. II). The corrected VOT were entered in a three-way (weeks \times voicing \times place) repeated-measures ANOVA for each subject; each cell contained VOTs arising from three repetitions of the /Cad/ syllable. t -tests for matched pairs contrasted corrected VOT in the two preactivation sessions pooled (called baseline sessions) with the last two post-activation sessions pooled (called final sessions).

II. RESULTS

A. Relation between syllable duration and VOT

Figure 1 shows the relation of VOT to syllable duration. Each panel shows the relation for one of the four implant users between the VOTs produced and the durations of the syllables in which they occurred. For comparison, the squares plot mean VOT in each class interval of syllable duration obtained from three hearing speakers asked to read lists of six CV syllables at widely varying rates (Volaitis and Miller, 1992). Mean VOTs of syllables beginning with voiced plosives are connected by solid lines, voiceless by dotted lines. Circles plot the corresponding measures from

the present study obtained in baseline (open circles) and final sessions (filled). (Data points based on $N=2$ or less are not plotted.)

Figure 1 shows that mean VOT varies with syllable duration—in speakers with normal hearing (squares) as well as in our deafened speakers both before (open circles) and after (filled circles) activation of their implant processors. With two exceptions, straight lines fit to the plots with the method of least squares have positive slopes; the exceptions correspond to FA's voiced plosives pre-activation and MC's voiced plosives post-activation; the slope in both these exceptions is negative and shallow, -0.04 . The measures obtained with normally hearing subjects are well fit by straight lines with $r[24]=0.78$ for the voiced plosives and $r[24]=0.93$ for the voiceless. A linear relation also appears to hold generally for our deafened speakers. Of the 16 plots of VOT versus syllable duration that in principle could be tested for linearity (2 conditions of voicing \times 2 processor states \times 4 speakers), the small ranges of syllable durations in this study precluded tests of linearity in seven cases; the median correlation coefficient in the remaining nine was $r=0.91$. The VOTs of voiceless tokens show a much larger effect of syllable duration than do those of the voiced tokens and this may reflect different underlying mechanisms in the production of VOT.

It will be observed that one effect of processor activation in the present study was to reduce syllable durations uttered by three of the four speakers (FA, FB, and FC); that is, filled circles frequently lie to the left of corresponding open circles in Fig. 1. The average reduction in syllable duration for the four speakers combined was 10.5%. All speakers significantly changed syllable durations over weeks (FA: $F[10,20]=4.8$, $p<0.01$; FB: $F[7,14]=28.5$, $p<0.01$; FC: $F[7,14]=9.2$, $p<0.01$; MC: $F[7,14]=3.5$, $p<0.01$) and all uttered longer syllables on the average when the syllables began with a voiceless as opposed to a voiced plosive (FA: $F[1,2]=244$, $p<0.01$; FB: $F[1,2]=374$, $p<0.01$; FC: $F[1,2]=14,400$, $p<0.01$; MC: $F[1,2]=301$, $p<0.01$). The three female speakers, but not MC, significantly reduced average syllable duration from baseline to final sessions 88, 61, and 27 ms, respectively (FA: $t[34]=13.2$, $p<0.01$; FB: $t[35]=12.6$, $p<0.01$; FC: $t[35]=4.8$, $p<0.01$).

Since syllable durations become shorter with activation of the cochlear implant, measures of VOT following processor activation potentially confound effects of two distinct factors: a general "postural" change in speaking rate and a selective change in implementing the voicing contrast. For example, because of the link between VOT and syllable duration, the observed drop in syllable duration following processor activation could reduce or cancel an increase in VOT caused by the speaker's newfound ability to hear the voicing contrast.

In light of these findings, VOTs were corrected for changes in syllable duration as follows. The mean syllable durations of voiced and of voiceless tokens in the first baseline session were computed separately for each speaker. Then, for each of a speaker's voiced and voiceless tokens in all the other sessions, the appropriate mean from the first baseline session (voiced or voiceless) was subtracted from the syllable duration of the token to obtain its change in

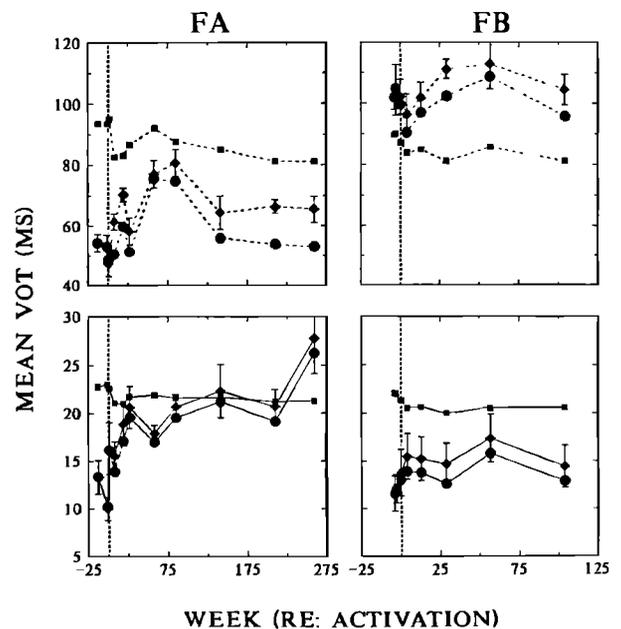


FIG. 2. The effects of processor use on VOT for speakers FA and FB. Mean VOT (circles) and mean corrected VOT (diamonds) are plotted as a function of weeks since processor activation (vertical line). Each point is the mean of nine determinations, three at each of three places of articulation. Vertical bars represent one standard error of the mean. The upper panels plot results for the voiceless plosives, the lower for the voiced, with a difference in scale: The upper VOT range is 80 ms, the lower 25 ms. The squares show values derived from data obtained from speakers with normal hearing by Volaitis and Miller (1992).

duration from baseline (generally, a reduction). The change in VOT associated with this change in syllable duration was obtained from the Volaitis and Miller data in Fig. 1, multiplying the syllable duration by the appropriate slope constant, 0.021 for voiced tokens and 0.132 for voiceless tokens. (These are the slopes of the straight lines fit by the method of least squares to the functions relating VOT to syllable duration.)² Finally, this change in VOT (generally, a decrease) was subtracted from the VOT of the token, to obtain (generally a higher) VOTc. VOTc was adopted as the dependent variable for examining the effects of activation of the cochlear implant on implementing the voicing contrast. Changes in VOTc following processor activation should be attributable largely to factors other than speaking rate.

B. Effects of processor activation on voicing

The effects of processor use on VOT are shown in Figs. 2 (for FA and FB) and 3 (for FC and MC). Mean VOTc (diamonds) and VOT (circles) are plotted as a function of weeks since processor activation (vertical line). The upper panels plot results for the voiceless plosives, the lower for the voiced, with a difference in scale: The upper VOT range is 80 ms, the lower 25 ms. The squares show values of mean VOT that would be expected from a group of hearing speakers who had the same average syllable duration as the implant user in each session. These hearing values were obtained as follows. The implant user's mean syllable duration in each session was computed for voiced and voiceless plosives separately. Then the value of VOT associated with that syllable duration in the

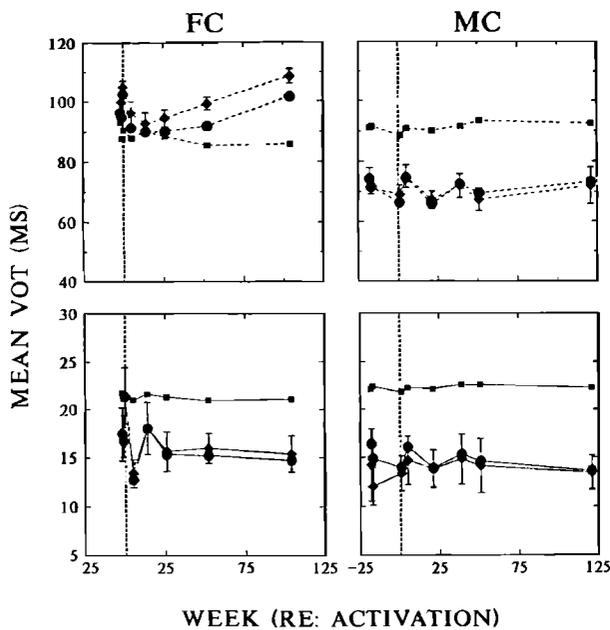


FIG. 3. The effects of processor use on VOT for speakers FC and MC. Legend as in Fig. 2.

Volaitis and Miller data shown in Fig. 1 was calculated using the regression equations for those data. (Values of standard error in the Volaitis and Miller data are too small to depict: They were approximately constant up to 500 ms at 1 ms for voiced plosive VOT and 2 ms for voiceless.) Although the values of VOT for hearing speakers in Figs. 2 and 3 are derived from data obtained with magnitude productions of speaking rate, Waldstein (1989) recorded five hearing speakers reading CVC(C) word lists and reported results that are similar. The average value of VOT for hearing speakers plotted in Figs. 2 and 3 is 21 ms for the voiced plosives, 88 ms for the voiceless; Waldstein's results are 16 and 92 ms, respectively. (Syllable durations were not reported in the latter study, so an exact comparison between the Volaitis and Miller and the Waldstein findings cannot be made here).

Three of four deafened speakers changed mean VOTc reliably over time following activation of their implant processors (FA: $F[10,20]=11.1$, $p<0.01$; FB: $F[7,14]=2.4$, $p=0.07$; FC: $F[7,14]=5.5$, $p<0.01$; MC: $F[7,14]=6.7$, $p<0.01$); the F ratio for speaker FB is not significant). Mean values of VOTc differed reliably depending on voicing (FA: $F[1,2]=790$, $p<0.01$; FB: $F[1,2]=1453$, $p<0.01$; FC: $F[1,2]=2678$, $p<0.01$; MC: $F[1,2]=2252$, $p<0.01$); voiceless plosives have longer VOTc. The difference between voiceless and voiced tokens in VOTc did not remain constant: There were significant interactions between voicing and weeks for FA and FC (FA: $F[10,20]=4.9$, $p<0.01$; FC: $F[7,14]=4.2$, $p<0.05$). (Place of articulation is discussed below.)

FA had quite short VOTc in both voiced and voiceless plosives pre-implant. Her plosives had roughly half the VOT of the group of hearing speakers. Following activation, both voiced and voiceless VOTc increased. After approximately 2 years of processor use, however, FA's voiceless VOTc fell some 15 ms. Whereas the difference between voiced and

voiceless mean VOTc had risen from about 40 ms preactivation to 60 ms within a year, it then fell 15 ms and remained at this level of voicing contrast for another three years. Her average VOTc in the final sessions was significantly higher than in the baseline sessions for both the voiced plosives and the voiceless ($t[17]=6.0$, $p<0.01$; $t[17]=4.3$, $p<0.01$). All three places of articulation (pooling over voicing) participated in the increase in VOTc for FA ($t[11]=4.4$, $p<0.01$; $t[10]=3.4$, $p<0.01$; $t[11]=4.4$, $p<0.01$) for bilabial, alveolar and velar plosives, respectively).

FB, like FA, had low values of VOTc for voiced plosives prior to the restoration of some hearing with the cochlear implant; her voiceless plosive means, however, were somewhat higher than the values for hearing speakers. Following processor activation, only voiced VOTc rose reliably ($t[17]=4.2$, $p<0.01$); voiceless VOTc fluctuated around a central value of 100 ms, well above values from hearing speakers. With voiced VOTc below and voiceless VOTc above normative values throughout the experiment, this speaker's implementation of the voicing contrast remains exaggerated at some 90 ms. Among the three places of articulation (pooling over voicing), only FB's alveolar plosives showed a reliable change in VOTc ($t[11]=2.35$, $p<0.05$).

Before activation of the implant speech processor, FC, like FB, uttered voiceless plosives with mean VOTc not far from those of speakers with normal hearing; the same is true of her voiced plosives. In the first session post-activation, recorded two days after the speech processor of her implant was turned on, VOTc increased about 5 ms. By the next session, five weeks later, voiced and voiceless VOTc fell slightly; however, voiceless VOTc climbed from the fifth to the eighth recording, terminating at a significantly higher level than pre-implant ($t[17]=2.2$, $p<0.05$); there is no reliable change in voiced VOTc. With voiceless VOTc increasing and voiced VOTc showing no net change, the separation of the two increased roughly 10 ms from baseline to final sessions ($t[17]=3.4$, $p<0.01$). Among the three places of articulation (pooling over voicing), only bilabial plosives increased VOTc reliably from baseline to final sessions ($t[11]=2.7$, $p<0.05$).

Subject MC showed no significant change in mean VOTc with processor activation and maintained voiced and voiceless VOTc well short of the measures from speakers with normal hearing. Approximately 60 ms generally separated VOTc means for his voiced and voiceless plosives; this was roughly the separation in the hearing group. MC differentiated voicing using VOTc to a degree intermediate between FA (40 ms in final sessions) and the other two deaf speakers (90 ms). For MC, one of the three places of articulation yielded reliable changes in VOTc.³

C. Effect of processor activation on place differences in VOT

The variation in VOTc over the three places of articulation (pooling voiced and voiceless tokens) was statistically significant for all four speakers (FA: $F[2,4]=24.6$, $p<0.01$; FB: $F[2,4]=55.4$, $p<0.01$; FC: $F[2,4]=10.5$, $p<0.05$; MC: $F[2,4]=97.2$, $p<0.01$) but small; an average of 6 ms separated adjacent places. These separations did not change sig-

TABLE II. Percent plosive identification errors (corrected for guessing post-activation) in assigned voicing or place. Each of four implanted subjects was tested at the number of weeks shown postactivation of her (his) speech processor.

| Week | FA | | FB | | FC | | MC | | | | |
|------|-------|-------|------|-------|-------|------|-------|-------|----|----|----|
| | Voice | Place | Week | Voice | Place | Week | Voice | Place | | | |
| 9 | 68 | 54 | 28 | 6 | 24 | 8 | 12 | 66 | 61 | 10 | 57 |
| 88 | 54 | 21 | 61 | 8 | 19 | 16 | 14 | 64 | | | |
| 140 | 18 | 49 | 65 | 8 | 6 | 20 | 14 | 54 | | | |
| 144 | 34 | 39 | | | | 28 | 12 | 58 | | | |
| | | | | | | 51 | 2 | 34 | | | |

nificantly for any subject from pre- to post-activation of the implant processor; the results are therefore not plotted. Apparent departures from the normative ordering of VOTc with place (front-mid-back) were observed with FA (both voiced and voiceless pre-activation, voiceless post), FC (voiceless pre and post), and MC (voiceless pre and post); the small differences in VOTc among adjacent places of articulation combined with the small sample sizes precludes any conclusion about these orderings. Of the 16 orderings measured (2 sessions×2 values of voicing×4 speakers) 11 were in the normative order in the baseline sessions and nine were in that order in the final sessions.

D. Speech reception tests

Table II presents the errors of voicing and of place in the subjects' plosive identifications. FA made 68% voicing errors (corrected for guessing)⁴ in a listening test administered soon after the speech processor of her implant was activated. Her error rate fell over subsequent months, and in the most recent administrations (at 140 and 144 weeks post-activation) she averaged 26% voicing errors. There is no discernible trend in this speaker's place errors, which remain high. FB seldom made voicing errors in identification, remaining at about 8% over a year post-activation. Her place errors fell by more than half to the low level of her voicing errors—6%—over this period. FC likewise makes few voicing errors; 1 year post-activation, the rate was at an all time low of 2%. Her place errors also appear to have fallen but remain high at 34%. Subject MC received the test only once, a little over a year post-activation. Like FB and FC, he made few voicing errors; his place errors were much more numerous, as observed with FA and FC.

III. DISCUSSION

A. Effect of processor activation on the voicing contrast

Before activation of their implant processors, all four speakers had VOTc averages in syllable-initial plosives that were shorter than normal or close to normal. In addition, implant user FA displayed less than normal difference in average VOTc of voiced and voiceless cognates pre-activation. Waldstein (1989, 1990) found related results with seven postlingually deafened adults. All but one of her speakers uttered voiceless plosives with shorter than normal VOT and showed less than normal differences in VOT between voiced

and voiceless cognates. The one exception was a speaker deafened at age 40 whose productions were within the range of a group of speakers with normal hearing also tested. The speakers in Waldstein's study showing the most disturbance of the voicing contrast were the two deafened before their teens. Since, as we have seen, prolonged profound deafness causes speakers to speak more slowly and VOT is accordingly lengthened, Waldstein's speakers' undershoot of normative values would probably be even more marked if a correction were applied to her VOT measurements to offset the component due to her speakers' low rates of speech.

Similarly, the teenage patient of Tartter *et al.* (1989) had abnormally long syllable durations that shortened (monosyllables and spondees) after the activation of her implant speech processor. VOT measured without correction for the change in rate was reported not to have changed post-activation; however, VOTc may have changed. This patient's production of another cue associated with the voicing feature, the starting frequency of the first-formant transition, did change significantly between pre-implant baseline and 1-year post-activation of her speech processor. Economou *et al.* (1992) also found no systematic change in VOT in a child deafened at age 6, implanted with a single-channel device at age 7, and reimplanted with the Nucleus-22 prosthesis at age 10 (although voiced stops did shorten VOT significantly in one of the recording sessions).

Three of the four speakers in the present study increased VOTc significantly following activation of their implant processors. Information about the speakers' abilities to discriminate voicing with their implants (Table II) obviously speaks directly to the issue of the sources of their VOTc changes shown in Figs. 2 and 3.

In the three years following activation of her speech processor, FA showed some improvement in her ability to correctly identify the voicing of plosives; her error rate at the last testing (correcting for guessing) was 34%. The other speakers infrequently confused voiced and voiceless plosives by ear at any time during the course of the study; their mean error rates averaged about 10% in repeated testing at variable intervals. Rabinowitz *et al.* (1992) reviewed a body of literature showing that temporal "envelope" cues are more readily available to implant users than spectral cues. The amplitude envelope for a voiceless stop has a prominent peak at the release burst and a long interval with low amplitude before vowel onset; that for the voiced stop has a less prominent peak and a shorter interval. Consistent with the hypothesis

that our patients were using temporal envelope cues in speech reception, Table II shows that, in their last test, they made 14% errors (corrected for guessing) in identifying the voicing of plosives but 34% error in identifying their place of articulation (however, FB's final error rate for place perception was quite low). All four subjects, then, presumably had access to the way in which others and they themselves contrasted voiced and voiceless segments.

Three of the speakers, FA, FB, and FC, seem to have used that access to change VOT. The upward trend in the VOTc of FA's voiced and voiceless plosives is roughly concurrent with an increasing trend in her discrimination of voicing. Likewise, the increase in mean VOTc of FB's voiced plosives is concurrent with her newly recovered ability to discriminate plosive voicing. Her voiceless VOTc does not increase, presumably because it is already somewhat above normal. Under the same conditions, however, FC's voiceless plosives do show increased VOTc.

Despite this evidence indicating an effect of processor activation on VOTc, activation did not inevitably lead to normalization of previously abnormal VOTc. FA improved on the listening test in the weeks following processor activation, but her final rate of voicing errors in identification was the highest of the four subjects and she also seems to have stabilized voiceless VOTc in final sessions at values well below normal. Speaker FC's voiced VOTc was somewhat shorter than normal pre-implant and remained so post-activation; it was, however, not as short as the means obtained from FA and FB, which did increase. Finally, although MC accurately perceived voicing using his cochlear implant, voiced and voiceless VOTc remained reduced in his speech after nearly three years of implant use. Thus the increases in VOTc in three of the four speakers following implant use are concurrent with evidence that these speakers can discriminate voicing aurally using their implants; however, reasonably accurate perception of plosive voicing did not always yield accurate production.

B. Effect of processor activation on place differences in VOT

Several studies have found that speakers with normal hearing increase VOT from bilabial to alveolar and from alveolar to velar places of articulation, for both voiced and voiceless English plosives. A study by Ohde (1984) found a significant effect of place of articulation on VOT, which rose systematically for the voiced plosives and from labial to alveolar for the voiceless; however, alveolar and velar voiceless plosive means were similar. The average change in VOT due to place was approximately 8 ms obtained from five hearing speakers reading CVC syllables embedded in a carrier phrase. Waldstein (1990) found a similar pattern of results with a mean place difference of 5 ms obtained from seven hearing speakers reading monosyllabic English words (however, there were numerous tokens with prevoicing). Volaitis and Miller (1992) obtained VOT differences between adjacent places of articulation averaging approximately 15 ms (this value was interpolated at a mean syllable duration of 350 ms, the average syllable duration of our four speakers in their final sessions); they had three speakers uttering runs of

six /Ci/ syllables at eight rates. Speakers in the present study frequently failed to differentiate plosive VOTc with respect to place of articulation both before and after processor activation. However, their mean VOTc difference between adjacent places of articulation was 6 ms overall, as reported above, and 8 ms in their final sessions. A strict comparison with the preceding studies is not appropriate since they used different elicitation materials and tasks and report measures of VOT not adjusted for speaking rate. Nevertheless, it appears that the place effect is to be found in the speech of implant users.

Kluender (1991) has suggested that this effect "is the result of articulatory constraints, likely being related to inertial constraints upon the articulators involved" (p. 84). To the extent that place differences in VOT are determined solely by the articulators involved, one would expect our implanted speakers to have differentiated place in the way that hearing speakers do, since presumably they use the same articulators in producing each of the plosives that speakers with normal hearing do. On the other hand, to the extent that place differences in VOT are determined by the way in which articulators are controlled to release the constriction, our speakers' anomalous place differences in VOT, where they occur, may reflect idiosyncratic ways, acquired during prolonged deafness, of articulating each of the plosives.

IV. SUMMARY AND CONCLUSIONS

Returning to the questions posed at the outset of this study, we find that all four of our deafened speakers presented some anomaly in the production of VOT in syllable-initial plosives. This outcome supports the hypothesis that hearing normally plays a role in maintaining the implementation of the voicing opposition.

With activation of their implant processors, three of four patients changed VOTc means in the direction of values obtained from speakers who had normal hearing. These patients also discriminated the voicing feature using their implants, supporting the hypothesis that one of the roles of hearing in normal speech production is to regulate the voicing contrast.

In a discussion of the effects on speech parameters of prolonged deafness and of the reintroduction of some hearing with cochlear implants, Perkell *et al.* (1992) present the hypothesis that hearing does not regulate speech production moment-to-moment. Instead: "(1) Self-hearing helps to calibrate production mechanisms by monitoring relations between the speaker's own articulations and his/her acoustic output. This calibration is performed in the face of numerous perturbations, ranging from adjustments in the speaker's body posture to changing listener demands. (2) The speaker can also validate his acoustic output by observing the behavior of his listeners and by detecting discrepancies between his own speech and theirs" (p. 2962).

Despite late-onset profound deafness of many years duration, all of our speakers maintained a distinct difference in the VOTc of their voiced and voiceless plosives before receiving their implants. This finding is consistent with the idea that they had a robust internal model of the relation between articulatory commands and the desired phonetic result, a model that continued to serve them when hearing was

lost. Three of the four speakers increased VOTc gradually when some hearing was restored. This finding of a gradual change in VOTc is consistent with the idea that the model guiding articulation is calibrated using auditory information when it is available. However, evidence that deafened adults recalibrate articulatory models after extended processor use when reading a corpus in the laboratory does not confirm, of course, that speakers with normal hearing rely on auditory information to validate articulatory programs for speech.

Reviewing a wide range of correlated changes in vowel articulation observed in their postlingually deafened speakers after activation of their implant processors, Perkell *et al.* (1992) conclude that almost all the changes they observed “were due to changes in the postural settings of physiological mechanisms regulating speaking rate, F_0 and SPL” (p. 2973); there was only modest evidence that patients used spectral information delivered by the prosthesis in changing patterns of vowel articulation. In the present study, VOT changes were found that appear to have been driven not only by a “postural” resetting of rate, but also to some extent by a fine-grain “tuning” of VOT itself. We hypothesize that this “tuning” was due to a recalibration of the mechanism controlling VOT after some self-hearing was restored. That recalibration may have been accomplished using predominantly temporal information from the speech wave.

The present findings do not reveal whether speakers were influenced by hearing the VOT separation in other people’s speech, in their own speech, or both. Further, this study cannot rule out the possibility that the change in VOT is brought about indirectly through the mediation of another speech production mechanism also affected by processor activation. The laryngeal parameters that control whether vibration occurs or not (Stevens, 1977) include the spacing between the folds, their stiffness and mass, and the pressure across the glottis. Three of the four speakers studied by Perkell *et al.* (1992), including FA and FB who served in the present study as well, showed increases in an indirect measure of glottal aperture following processor activation; glottal spreading delays the onset of voicing. FA and FB also markedly lowered fundamental frequency, consistent with a slackening of the vocal folds and a reduction of VOT. Investigations in progress in our laboratory that add measures of air flow and intra-oral air pressure to all of the preceding measures of plosive production in patients with implants should help to clarify whether factors other than the discrimination of VOT contributed to the changes in VOTc observed in this study.

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, and for his helpful remarks on the manuscript for this article. Dr. Joanne Miller, Department of Psychology, Northeastern University, kindly gave us the measures of normally hearing speakers’ VOTs and syllable durations obtained during a magnitude production task, and discussed them with us. We appreciate the helpful discussion of this research with the Language and

Cognition Group of the Department of Psychology, Northeastern University and with Dr. Melanie Matthies and Dr. Mario Svirsky of the Research Laboratory of Electronics, Massachusetts Institute of Technology. Dr. Donald Eddington, Massachusetts Eye and Ear Infirmary, kindly provided us with the results from the administration of the MAC Battery. Helpful comments on an earlier draft of this article came from Dr. Vivien Tartter, City College, CUNY, New York. 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, Principal Investigator.

¹To facilitate subject identification in this paper, they are called FA, FB, FC, and MC (for the three females and the male speakers, respectively). For cross reference to other publications from the Massachusetts Eye and Ear Infirmary that include data from these subjects, they are S09, S15, S23, and S19, respectively. The same scheme was used in Perkell *et al.* (1992), which reported on changes in vowel production in four implanted speakers including FA and FB. MA and MB from that study were not included in the present one because their deafness dated from early childhood.

²Volaitis and Miller (1992) report VOTs for syllable durations varying between 150 and 750 ms. Straight lines were also fit to the subset of those durations between 250 and 500 ms—the range of syllable durations encountered in the present study. Those least-squares slopes were 0.166 for voiceless tokens and 0.025 for voiced. Applying these, slightly steeper, slopes to the VOT corrections, yielded VOTc means that were, on the average, 2.1% greater for voiceless utterances and 0.3% greater for voiced utterances in block three of the present experiment. The more conservative corrections, based on the entire data set from Miller and Volaitis, were employed.

³On two occasions prior to activation of his implant processor, MC replaced the intended voiced alveolar plosive with the homologous voiced affricate. These tokens were not scored and were represented in the repeated-measures ANOVA by the mean of the other two determinations for this plosive in the same session.

⁴The corrected error rate equaled: (1) [(percent correct – chance level)/(100 – chance level)].

Boothroyd, A., Hanin, L., and Medwetsky, L. (1988). “Speech production changes in cochlear implantees,” unpublished report, Speech and Hearing Sciences Research Center, City University of New York.

Cowie, R. I., and Douglas-Cowie, E. (1983). “Speech production in profound post-lingual deafness,” in *Hearing Science and Hearing Disorders*, edited by M. E. Lutman and M. P. Haggard (Academic, New York), pp. 183–231.

Davis, P. J., Bartlett, D., and Luschei, E. S. (1992). “Coordination of the respiratory and laryngeal systems in breathing and vocalization,” *Natl. Center Voice Speech Status Progr. Rep.* 2, 59–81.

Economou, A., Tartter, V. C., Chute, P. M. and Hellman, S. A. (1992). “Speech changes following reimplantation from a single-channel to a multichannel cochlear implant,” *J. Acoust. Soc. Am.* 92, 1310–1323.

Eddington, D. (1983). “Speech recognition in deaf subjects with multichannel intracochlear electrodes,” *Ann. N.Y. Acad. Sci.* 405, 241–258.

Gold, T. (1978). “Speech and hearing skills: A comparison between hard-of-hearing and deaf children,” unpublished doctoral dissertation, City University of New York.

Gold, T. (1980). “Speech production in hearing impaired children,” *J. Commun. Disord.* 13, 397–418.

Henke, W. (1989). *MITSYN Languages, Language Reference Manual* (Author, Belmont, MA).

Hutchinson, J., and Smith, L. (1976). “Aerodynamic functioning in consonant production by hearing-impaired adults,” *Audiol. Hear. Educ.* 2, 16–19.

Kluender, K. (1991). “Effects of first formant onset properties on voicing judgments result from processes not specific to humans,” *J. Acoust. Soc. Am.* 90, 83–96.

Lane, H. L., Perkell, J., Svirsky, M., and Webster, J. (1991). “Changes in

- speech breathing following cochlear implant in postlingually deafened adults," *J. Speech Hear. Res.* **34**, 526–533.
- Levitt, H., and Stromberg, H. (1983). "Segmental characteristics of the speech of hearing-impaired children: factors affecting intelligibility," in *Speech of the Hearing Impaired*, edited by I. Hochberg, H. Levitt, and M. J. Osberger (University Park, Baltimore, MD).
- Mahshie, J. J., and Conture, E. G. (1983). "Deaf speakers' laryngeal behavior," *J. Speech Hear. Res.* **26**, 550–559.
- Monsen, R. B. (1976). "The production of English stop consonants in the speech of deaf children," *J. Phonet.* **4**, 29–41.
- Ohde, R. N. (1984). "Fundamental frequency as an acoustic correlate of stop consonant voicing," *J. Acoust. Soc. Am.* **75**, 224–230.
- Owens, E., Kessler, D., Raggio, M., and Schubert, E. (1985). "Analysis and revision of the minimal auditory capabilities (MAC) battery," *Ear Hear.* **6**, 280–287.
- 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., 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.
- Rabinowitz, W. M., Eddington, D. K., Delhome, L. A., and Cuneo, P. A. (1992). "Relations among different measures of speech reception in subjects using a cochlear implant," *J. Acoust. Soc. Am.* **92**, 1869–1881.
- Smith, C. (1975). "Residual hearing and speech production in deaf children," *J. Speech Hear. Res.* **18**, 795–811.
- Stevens, K. (1977). "Physics of laryngeal behavior and larynx modes," *Phonetica* **34**, 264–279.
- Svirsky, M., Lane, H., Perkell, J., 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., 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. (1989). "Acoustic characteristics of the speech of the postlingually deafened: implications for the role of auditory feedback in speech production," unpublished doctoral dissertation, Brown University.
- Waldstein, R. (1990). "Effects of postlingual deafness on speech production: Implications for the role of auditory feedback," *J. Acoust. Soc. Am.* **88**, 2099–2114.
- Whitehead, R., and Barefoot, S. (1980). "Some aerodynamic characteristics of plosive consonants produced by hearing-impaired speakers," *Am. Ann. Deaf* **125**, 366–373.
- Youngblood, J., and Robinson, S. (1988). "Ineraid (Utah) multichannel cochlear implants," *Laryngoscope* **98**, 5–10.