

November 01, 2008

Perception and production of /r/ allophones improve with hearing from a cochlear implant

Melanie L. Matthies

Boston University; Massachusetts Institute of Technology

Frank H. Guenther

Boston University; Massachusetts Institute of Technology

Margaret Denny

Massachusetts Institute of Technology

Joseph S. Perkell

Massachusetts Institute of Technology; Boston University

Ellen Burton

Massachusetts Institute of Technology

See next page for additional authors

Recommended Citation

Matthies, Melanie L.; Guenther, Frank H.; Denny, Margaret; Perkell, Joseph S.; Burton, Ellen; Vick, Jennell; Lane, Harlan; Tiede, Mark; and Zandipour, Majid, "Perception and production of /r/ allophones improve with hearing from a cochlear implant" (2008). *Psychology Faculty Publications*. Paper 13. <http://hdl.handle.net/2047/d20000860>

Author(s)

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

Perception and production of /r/ allophones improve with hearing from a cochlear implant^{a)}

Melanie L. Matthies

Department of Speech Language and Hearing Sciences, Boston University, Boston, MA 02215 and Speech Communication Group, Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Frank H. Guenther

Department of Cognitive and Neural Systems, Boston University, Boston, MA 02215-0000 and Speech Communication Group, Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Margaret Denny

Speech Communication Group, Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Joseph S. Perkell

Speech Communication Group, Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139; Department of Cognitive and Neural Systems, Boston University, Boston, MA 02215; and Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Ellen Burton

Speech Communication Group, Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Jennell Vick

Department of Speech and Hearing Sciences, University of Washington, and Speech Communication Group, Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Harlan Lane

Department of Psychology, Northeastern University, Boston, MA and Speech Communication Group, Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Mark Tiede

Speech Communication Group, Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139 and Haskins Laboratories, New Haven, CT 06511

Majid Zandipour

Speech Communication Group, Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139 and Department of Cognitive and Neural Systems, Boston University, Boston, MA 02215

(Received 23 January 2008; revised 19 August 2008; accepted 20 August 2008)

Tongue shape can vary greatly for allophones of /r/ produced in different phonetic contexts but the primary acoustic cue used by listeners, lowered F3, remains stable. For the current study, it was hypothesized that auditory feedback maintains the speech motor control mechanisms that are constraining acoustic variability of F3 in /r/; thus the listener's percept remains /r/ despite the range of articulatory configurations employed by the speaker. Given the potential importance of auditory feedback, postlingually deafened speakers should show larger acoustic variation in /r/ allophones than hearing controls, and auditory feedback from a cochlear implant could reduce that variation over time. To test these hypotheses, measures were made of phoneme perception and of production of tokens containing /r/, stop consonants, and /r/+stop clusters in hearing controls and in eight postlingually deafened adults pre- and postimplant. Postimplant, seven of the eight implant speakers did not differ from the control mean. It was also found that implant users' production of stop and stop+/r/ blend improved with time but the measured acoustic contrast between these was still better in the control speakers than for the implant group even after the implant users had experienced a year of improved auditory feedback.

© 2008 Acoustical Society of America. [DOI: 10.1121/1.2987427]

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

Pages: 3191–3202

I. INTRODUCTION

Variability in the articulation of /r/ in American English has been studied extensively (see [Delattre and Freeman, 1968](#); [Stevens 1998](#); [Westbury et al., 1998](#); [Zhou et al., 2008](#)) with respect to the strong influence of adjoining phonemes and range of individual speaker productions. Large variations in vocal tract shape for /r/ can be observed in different phonetic contexts, but these routinely provide a lowered F3 in the acoustic signal which in turn supplies a reliable perceptual cue to the listener. In discussing the six major types of tongue shapes used in American English, [Delattre and Freeman \(1968\)](#) note, "...the spectrographic pattern is more significantly affected by syllabic position than by tongue shape" (page 52). Variety in vocal tract shape has not been found to distort the resultant acoustic output (for F3 and lower formants) for /r/ produced by normal speakers nor do listeners perceive differences between retroflexed and bunched articulations. [Zhou et al. \(2008\)](#) found significant differences between F4/F5 patterns in speakers with bunched versus retroflexed /r/ productions but concluded that these patterns affected the perception of speaker identity rather than phonetic content. With a range of articulatory configurations providing the lowered F3 that corresponds to the /r/ percept, a strong case can be made that the speaker's task is to provide a consistent acoustic cue across a variety of contexts and to implement whatever articulatory configuration will best lower F3 in a given context.

The observation that many articulatory configurations can be found in American English /r/ indicates significant complexity in the production of /r/. Articulatory variability and the influence of context may present difficulties in the acquisition of /r/ by children and second-language learners. In a clinical intervention study, [Adler-Bock et al. \(2007\)](#) described /r/ as one of the most complex phones in English because of the multiple articulatory configurations involved in its production. Auditory feedback is a potential mechanism for constraining the acoustic variability of /r/ so that the speaker can provide the listener with a stable cue for perception. The speech motor control system can use auditory feedback in planning the vocal tract shape so that the listener's percept remains /r/ despite a range of articulatory configurations employed by the speaker. The many-to-one relation between articulation and realized /r/ has not been studied extensively in postlingually deafened adults. This is an interesting population because the process of using auditory feedback in lowering F3 will be particularly challenging for a speaker who has difficulty hearing this cue. Consequently /r/ production may degrade in persons with hearing loss.

A. Effects of hearing loss and improved auditory feedback

While speakers with acquired deafness may remain highly intelligible for years after losing hearing, some deterioration in articulatory precision is to be expected (see [Lane et al. 1995](#)) and certain phonemes may be more susceptible

than others. All phonemes can show variation due to coarticulation, but the semivowels are particularly labile with changes in phonetic context. The literature suggests that the production of semivowels is not characterized by strong articulatory anchor points, saturation effects, or robust orosensory feedback. Speakers therefore may rely more on auditory information to guide the production of /r/ than for other consonants. Postlingually deaf speakers may have difficulty in consistently producing lowered F3 and be less able to manage the disparate /r/ articulations induced by varying context. In accord with this idea, some previous work has indicated that deafness may cause a decrease in the /r/-/l/ contrast ([Perkell et al., 2001](#)). Because of hearing loss, the auditory feedback of deaf speakers can be insufficient for accurately and reliably mapping between the many articulatory possibilities of /r/ due to context variation and the goal of lowering F3 to produce an intelligible percept for the listener (see [Guenther et al., 1998](#); [Guenther et al., 2006](#); [Perkell et al., 2001](#)).

When adventitiously deafened adults receive cochlear implants, they generally show improvements in speech production that are consistent with the objective of transmitting an intelligible message to the listener. Acoustic measures reveal diminished phoneme contrasts in the speech of those deafened adults compared to hearing controls. When some hearing is restored with an implant, gradual improvements are commonly seen (as measured via acoustic analyses and/or intelligibility testing of the speech). Improvements in speech from postlingually deafened adults have been demonstrated longitudinally over several months or even years. In experiments that switch the implant speech processor off and on, phoneme contrasts can improve quickly when auditory feedback is available. These findings apply both to vowel contrasts ([Perkell et al., 1992](#); [Lane et al., 2001](#); [Kishon-Rabin et al., 1999](#)) and consonant contrasts ([Lane et al., 1995](#); [Matthies et al., 1994, 1996](#); [Waldstein, 1990](#)).

B. A modeling framework

To provide a framework for studying the normal and disordered production of phonemes, models of speech production try to account for the many-to-one relationship between articulation and the acoustic output. In an acoustic and articulatory study of /r/ production in seven speakers with normal hearing, [Guenther et al. \(1999\)](#) observed that the participants demonstrated a trade-off between front-cavity length and the size and length of the constriction in different contexts, keeping F3 low—the primary cue for /r/. Using data obtained with an electromagnetic articulometer, [Guenther et al. \(1999\)](#) found that productions of /r/ sounds in varying phonetic contexts were acoustically similar despite large differences in their associated articulatory configurations. They noted further that speakers were modifying front-cavity length and constriction length to preserve a total cavity volume that would maintain the acoustic target (low F3) of /r/. A modeling study of articulatory and acoustic variability based on two of the speakers of [Guenther et al. \(1999\)](#) supported the idea that the control of /r/ production is based on a low F3 auditory/acoustic target ([Nieto-Castanon et al., 2005](#)).

^{a)} A portion of this work was presented at the 150th meeting of the Acoustical Society of America in Minneapolis, MN, Fall, 2005.

The claim for an auditory goal for /r/ also receives support from a study of native speakers of Japanese learning English as a second language. Such speakers typically have difficulty producing the English /r/-/l/ contrast. Bradlow *et al.* (1999) used a perceptual training paradigm that presented many exemplars to their participants to help them form new phonetic categories, without any production training. They found improvements in production that were correlated with improvements in perception. Along similar lines, Ingram and Park (1998) in a study with Japanese and Korean speakers noted that both pretraining phonological knowledge of the /r/-/l/ contrast and discriminability of the acoustic differences between /r/ and /l/ were important for the identification and discrimination of the phonemes.

C. Study objectives

For the current study, an experiment was designed to examine the role of auditory feedback in the production of /r/ by studying the effect of context on /r/ when produced by postlingually deafened adults. The current study is based on the general hypothesis that speakers attempt to satisfy listener demands for clarity while minimizing their expenditure of effort (Lindblom, 1990; Perkell *et al.* 1997, 2004). Speakers are presumed to have internal models of their speech production mechanisms and acquire acoustic-phonetic knowledge that combine to guide their efforts to produce intelligible speech. In this internal model, movement goals for phonemes can be specified in the articulatory and auditory spaces, as set forth in the Directions Into Velocities of Articulators (DIVA) neurocomputational model of speech motor planning (Guenther, 1995; Perkell *et al.*, 2001). These goals are acquired, maintained, and achieved through the use of feedback and feedforward control mechanisms, which consist of acquired mappings between articulatory and sensory parameters. In particular, there is a primary auditory goal for /r/, which consists mainly of a range of low values of F3. This single auditory goal region can be achieved with many different configurations of the tongue but two, bunched and retroflexed, represent the extremes of this continuum. According to the model, there would be two articulatory-to-auditory mappings for /r/, one for each extreme articulatory configuration. In postlingually deafened adults who have been without hearing for a long time, such mappings are presumed to have degraded somewhat, leading to more variability in the production of auditory/acoustic goals. These ideas led to the formulation of three hypotheses that were tested on two groups of speakers, postlingually deafened adults who received cochlear implants and normal-hearing controls.

The study tests a *context variability hypothesis*: varying phonetic context (as in the nonsense words wabrav, wadrav, wagrav, and warav) will yield more variable /r/ productions by implant users than by controls. The study also evaluates an *allophonic consistency hypothesis*: implant users whose speech perception has improved with experience using their implant will yield more consistent /r/ productions across phoneme context, similar to the pattern exhibited by hearing controls. According to the model, experience using their im-

plants would allow these speakers to recalibrate auditory feedback mechanisms and, using that recalibrated feedback, to modify mappings between disparate articulatory configurations and a more consistent auditory goal for /r/ across phonemes. An additional look at the role of auditory feedback was devised by manipulating the auditory feedback condition for the normal-hearing speakers with masking noise and for the cochlear-implant users by turning their speech processors off during part of the experimental sessions. The *consonant contrast hypothesis* tests that the contrast between stop+r and stop-alone productions is greater for speakers with normal hearing than for implant users. If auditory feedback is useful for maintaining precise articulation of clusters, these contrasts can be expected to vary with feedback state and experience with the cochlear implant.

II. METHOD

A. Participants

The experimental group consisted of six male and two female postlingually deaf, adult, paid volunteers, who received cochlear implants at an average age of 52. The implant was either the Clarion (Advanced Bionics, CIS strategy; Wilson *et al.* 1995) or the Nucleus device (Cochlear Corp.; Blamey *et al.* 1987; McKay and McDermott, 1993). The implant users were referred by the Massachusetts Eye and Ear Infirmary or the University of Massachusetts Memorial Medical Center. The control group consisted of six females and one male; they had no reported difficulties with speech or hearing. All were paid volunteers with an average age of 38, older than the standard college students and therefore are more similar in age to the cochlear-implant users. The implant users were tested three times: preimplant (before their speech processors were turned on), after one month of experience with the implant, and after a year of implant use. Most of the normal-hearing control speakers served for a single session, but several were tested twice to confirm that the intersession variability for these measures was low. Details about the study participants are given in Table I. Each speaker participated in tests of speech production and perception.

B. Speech perception test

For the allophonic consistency hypothesis to be tested, participants with hearing loss must have improved their speech perception through the use of their cochlear implants. Multiple repetitions of the consonants /b, d, g, p, t, k, s, ʃ, z, r, and l/ in a /Cad/ environment were recorded by a normal-hearing male speaker. These were presented at a comfortable listening level to the implant users preimplant and at one month and one-year postimplant to measure longitudinal changes in consonant perception. Implant candidates who wore hearing aids preimplant used their aids during the phoneme recognition task in the session prior to implant activation but in the sessions at one month and one-year postactivation, they used only their cochlear implants. Further methodological details and results of the speech perception task have been given elsewhere (Lane *et al.*, 2007).

TABLE I. Participant demographic information, vowel and consonant percent correct identification for implant users, and normal-hearing controls. (Asterisks denote missing data.)

Implant users		FI	FJ	MJ	MK	ML	MM	MO	MP
Age at onset and (duration) profound hearing loss		54 (2)	45 (1)	43 (6)	28 (0)	48 (0)	72 (6)	67 (5)	26 (10)
Cochlear implant		Nucleus 24	Clarion High-Focus	Clarion High-Focus	Clarion Auria	Nucleus 24	Clarion High-Focus	Clarion High-Focus	Clarion High-Focus
Vowel perception (% correct)	Pre	29	26	35	25	64	32	51	27
	1 year	62	55	96	95	66	77	*	74
Consonant perception (%)	Pre	26	25	21	24	21	32	49	21
	1 year	44	57	74	92	82	50	*	74
Normal-hearing controls		FNH1	FNH2	FNH3	FNH4	FNH5	FNH6	FNH7	MNH1
Vowel perception (% correct)		99	98	99	99	98	*	96	97
Consonant perception (%)		91	90	99	99	96	96	98	99

Participants identified the presented consonant, choosing from a full set of 11 /Cad/ options displayed on the computer monitor; they used a keyboard or a touch screen to indicate their choice. After a practice run, participants were presented with 12 repetitions of each token arranged in semirandom order, and the responses were recorded and stored for further analysis. The normal-hearing participants were tested once.

C. Speech production elicitation

For each session, the speakers were seated in a sound-attenuating room in a comfortable chair with a head-mounted electret microphone (Audio-Technica, model AT803B) placed at a fixed distance of 20 cm from the lips. Elicitation utterances were randomized prior to the testing session and presented via a computer monitor to prompt the speaker's productions. The microphone output was low pass filtered at 7.2 kHz and digitized in real time at a 16 kHz sampling rate. In each session and condition, speakers produced 15 repetitions of words containing /r/, wabrav, wadrav, wagrav, and warav and the contrasting words without /r/, wabav, wadav, and wagav. The words were embedded in the carrier phrase "Say__ for me" (see Boyce and Espy-Wilson, 1997; Guenther *et al.*, 1999).

In the experimental condition in which feedback was blocked (feedback off), participants with normal hearing received masking noise at 95 dB sound pressure level (SPL) through calibrated TDH-39 headphones. The noise was generated by a custom-built device (Technical Collaborative, Lexington, MA) and was approximately speech shaped, with a spectral envelope that rolled off at 6 dB per octave. The implant users kept the normal settings of their speech processors for the feedback-on conditions and turned off their processors for the feedback-off conditions. For the first half

of the experiment, speakers had full access to auditory feedback (feedback on). For the second half of the experiment, implant users turned off their speech processors while participants with normal hearing received speech-shaped noise during their productions (feedback off). With seven test words and 15 repetitions of each, there were 105 utterances in each of the two listening conditions. Test words were randomized within each condition.

1. Acoustic analysis

As schematized in Fig. 1, the first three formant frequencies were extracted algorithmically from a linear predictive coding (LPC) spectrum at the consonant-/r/ boundary (dotted vertical line) and at the midpoint of the following /a/. The LPC filter order was chosen to optimize formant extraction

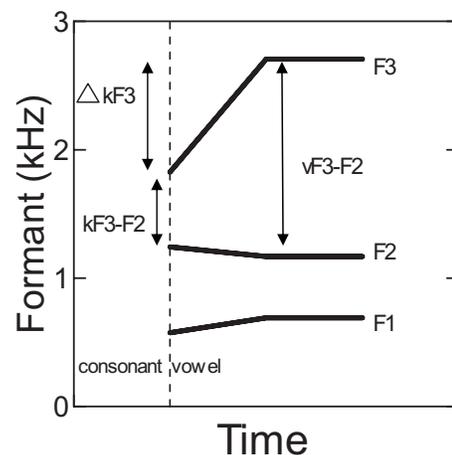


FIG. 1. Schematized formant tracks that depict measurements taken during the vowel, from the consonant-vowel boundary (dotted line) through the midvowel, and how these were used for analyses.

for each speaker. Initially, the extraction algorithm displays a broadband spectrogram with superimposed formant trajectories. If the three formants at the midvowel point were detected in frequency regions expected for /a/, the experimenter accepted that token with those values as complete. The analysis parameters (LPC order, window size, etc.) had gender-specific default settings, but the experimenter could adjust these parameters for individual speakers. If the algorithm did not succeed in detecting all of the necessary formants, the experimenter could vary the measurement time slightly. A small shift in the analysis time (to a different phase in the glottal cycle) would usually result in usable formant values. The detected formant values could also be modified manually if a formant peak was missing because of excessive bandwidth or peak merging. Such corrections were guided by observation of the concurrently displayed discrete Fourier transform (DFT) spectrum. In general, the articulation of /r/ can be classified as a lingua-palatal glide or semi-vowel (Ong and Stone, 1998). In accordance with this idea, the analysis extracted values from the formant trajectories according to the recognition algorithm of Espy-Wilson (1992) and theory described in Stevens (1998).

Six measurements were obtained from the formant tracks: F1, F2, and F3 at the consonant-vowel boundary (vertical dotted line) and F1, F2, and F3 at the midvowel. Analyzed variables included F1 and F2 at the midvowel, the difference between F2 and F3 at the consonant-vowel boundary ($kF3-F2$) and at the midvowel ($vF3-F2$), and the change in F3 ($\Delta kF3$) through the vowel. The arrow labeled $\Delta kF3$ in Fig. 1 designates the change in F3 (i.e., the F3 transition) from its minimum at or near the consonant-vowel boundary (dotted vertical line) to the middle of the vowel. The consonant-vowel boundary alone was used for the singleton consonant tokens (wabav, wadav, and wagav) without regard to whether F3 had reached a minimum (or maximum) but the minimum F3 was used for all /r/ tokens. Thus the extraction and analysis of the acoustic data were as parallel as possible across token types. The midvowel formants for /a/ were expected to be fairly similar across contexts while the formant transitions from the consonant-vowel boundary were expected to be strongly influenced by the specific consonant environment. In general, /r/ is characterized by a small initial difference between F3 and F2 ($kF3-F2$) with a sharp rise in F3 by midvowel ($vF3-F2$). McGowan *et al.* (2004) successfully used these formant-based measures in their study of /r/ production and found them to be robust across extensive morphological differences among their study participants, who were as young as 14 months of age.

2. Estimating variability of /r/ relative to /a/

To illustrate the variability phenomenon of interest in /r/ production, consider the disparate frequency distributions of the excursion of F3 transitions for /gr/ and /dr/ productions from a single speaker with hearing loss in the left half of Fig. 2. The distributions of the acoustic measure from a retroflexed /r/ produced in wadav compared to those from a bunched /r/ produced in wagrav are expected to be distinct from one another if the speaker did not have well-tuned map-

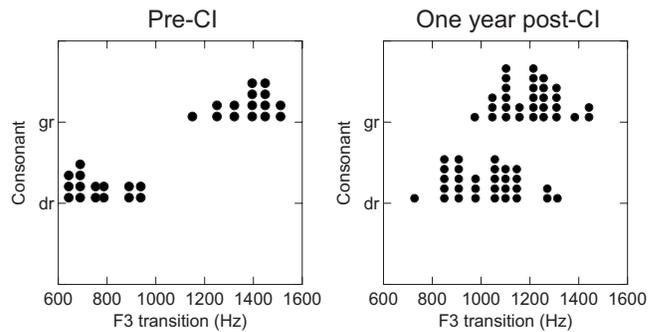


FIG. 2. A comparison of F3 transition data for wadav (likely to be retroflexed) and wagrav (likely to be bunched) as a function of experience with the cochlear implant for participant ML.

pings between the two different articulatory configurations and a single auditory goal for /r/. Indeed this was clearly the case for the implant user whose data are shown in Fig. 2 when acoustic parameters of /r/ were measured preimplant. In this speaker, considerable allophonic variation between /dr/ and /gr/ productions was seen in the preimplant productions (left panel); however, the distribution of the F3 transitions showed a postimplant pattern (right panel) consistent with normal hearing. That is, the distributions largely overlap in the one-year session (which includes both feedback conditions, right-hand panel), indicating less contextual influence on the acoustic realization of /r/. Other speakers showed this increase in allophonic overlap for the F3 transition but to a more limited extent, possibly due to the range and variety of articulation strategies that speakers use to produce /r/ (see Guenther *et al.*, 1999).

To provide a metric that was robust across all speakers, a ratio of /r/ dispersion to /a/ dispersion was constructed. As shown in Fig. 3, variability for /a/ was estimated as the dispersion of midvowel values of F1 and F2 for individual tokens around the center of their distribution (shown by arrows in the figure) in the F1, F2 plane. Dispersion was calculated as the Euclidean distance of each token from the distribution mean, averaged across all wabav, wadav, and wagav tokens. Dispersion was calculated for /r/ in a parallel way using values in the plane defined by the F3 transition (in hertz) and the F3-F2 difference (in hertz) at the /r/-vowel boundary.

A variability index was defined as the ratio of /r/ dispersion (average of distances to the /r/ center) to /a/ dispersion (distances to the vowel center). To calculate the variability index, tokens were matched pairwise by repetition number for wadav to wabav, wadav to wadav, and wagrav to wagrav. Pairing was used to provide approximate comparability for the dispersion comparisons with respect to potential warm-up and fatigue effects across the length of the experiment. For each pair, the /r/ distance to center was divided by the /a/ distance to the center. The variability index is the average of these ratios across multiple token pairs; it is closer to 1.0 when the dispersion measure for /r/ was more similar to the dispersion for /a/ as found in the bottom panel. When /r/ variability is much greater than vowel variability, the variability index value is much greater than 1.0, as seen in the top panel of Fig. 3. When the allophonic variability is reduced, /r/ dispersion also decreases while the /a/-dispersion

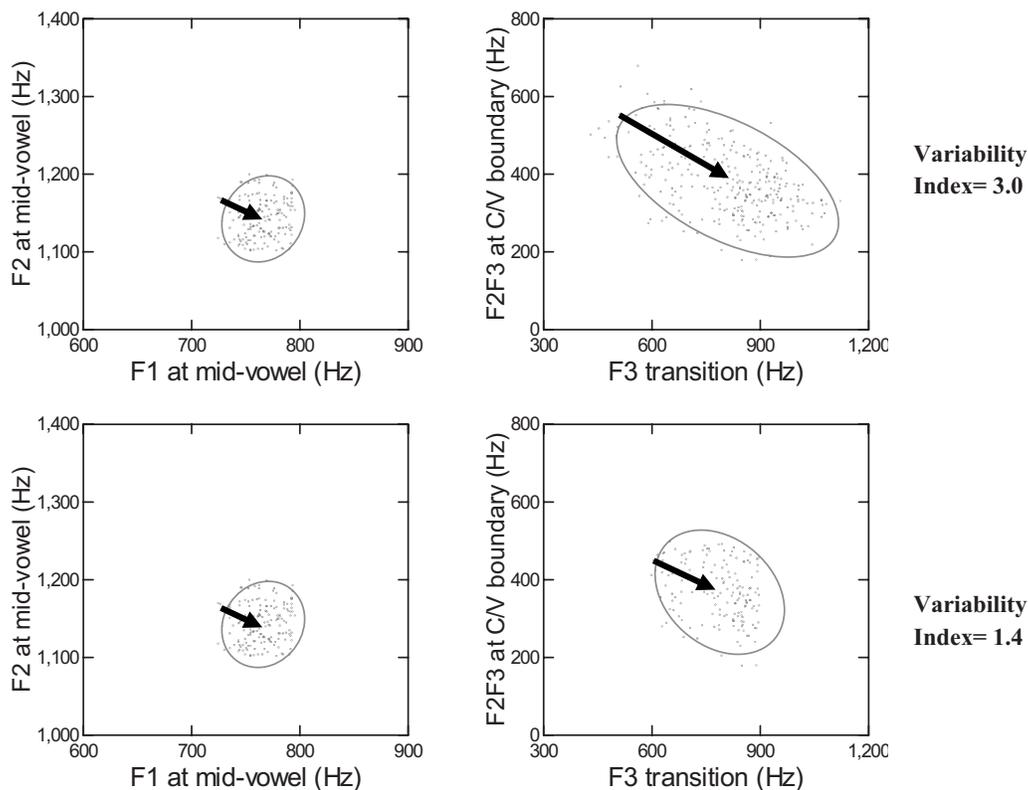


FIG. 3. Schematized comparison of a small variability index (lower panels) with a larger variability index (upper panels) due to greater /r/ dispersion in the upper right-hand plot. The /r/ dispersion is plotted in the domain of the F3–F2 difference at the beginning of the vowel by the F3 transition from vowel beginning to midvowel. The vowel dispersion is plotted in the F1, F2 domain.

remains roughly constant, yielding smaller variability indices (lower panels in Fig. 3). Log-based dispersion metrics were found to have a more normal distribution than linear ones, so logarithmically transformed measures were used for statistical analyses.

III. RESULTS

A. Speech perception

Vowel- and consonant-identification results for the normal-hearing participants are presented in the last two rows of Table I; they show very high or perfect scores. Table I also illustrates considerable gains in vowel and overall consonant identification for the implant users from the preimplant condition to one-year postimplant. As shown in Fig. 4, the implant users' overall proportion correct for consonant identification rises from less than 30% in the pre-cochlear implant (CI) condition to nearly 70% at the one-year session. Overall consonant identification and identification of the semivowels, voiced stops, and sibilants progress most rapidly between the preimplant condition and one-month testing period. Paired *t*-tests for /r/, sibilants, and /b, d, g/ identification revealed significant differences between pre- and one-month results for all three consonant types but no statistically reliable differences between the one month and the one-year sessions. The perceptual improvements noted in Table I and Fig. 4 support the goal of testing the allophonic consistency hypothesis to determine if corresponding speech production benefits could be observed when comparing preimplant to postimplant speech.

B. Speech production

The speech production results are presented in three sections. The first section reports findings from between-group comparisons to assess the context variability hypothesis, which is concerned with the variability of /r/ allophones for the implant group as compared to the speakers with normal hearing. The second section presents longitudinal results to evaluate the allophonic consistency hypothesis, which predicts that implant users will show a reduction over time in /r/ variability across phoneme context, making their results approach those from controls. The third section compares con-

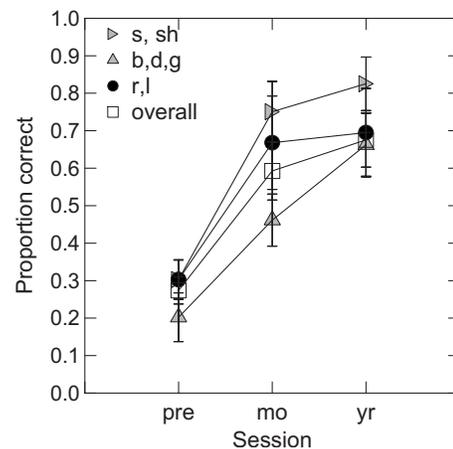


FIG. 4. Proportion correct from 11-item consonant confusion matrices constructed from results of consonant identification (12 repetitions per consonant) at a comfortable listening level.

trasts implemented between singleton stops and stop+/r/ in the two groups, thereby addressing the consonant contrast hypothesis, which states that those contrasts will be greater for speakers with normal hearing. The effects of experimentally induced auditory feedback changes (by adding masking noise) are included to provide a condition with greater difficulty (and therefore more variability) in the results for the normal-hearing speakers which may in turn provide a more realistic comparison for the results from the cochlear-implant users.

1. Between-group comparisons for variability analyses

A variability index was defined for each speaker as the average ratio of /r/ dispersion (average distance to the /r/ distribution center) to /a/ dispersion (average distance from their vowel mean), as explained in Sec. II. This ratio, rather than /r/ variability alone, was used to reduce between-speaker variability, particularly the potential effect of gender. Group averages for the implant users and normal-hearing participants were calculated by pooling values from individual token contrastive pairs across all the participants in each group. Although measures such as F0 and the formants in these words clearly varied with gender, the frequency distributions of variability indices for the two female implant users were statistically indistinguishable from those of the male participants.

The mean variability index for the control speakers with normal hearing was 1.22 (sd=0.35); there was greater variability in /r/ dispersion than in /a/ dispersion. When auditory feedback was disrupted with masking noise (feedback off) in normal-hearing controls, larger variability indices (mean = 1.33, sd=0.40) were found ($F=13.9$, $p<0.001$, and effect size $\eta p^2=0.12$). The main effect for plosive place of articulation was also significant ($F=14.5$, $p<0.001$, $\eta p^2=0.13$) with /br/ comparisons to /a/ in the feedback-on condition having the smallest variability index (mean=1.08) and /gr/ comparisons to /a/ having the largest (1.35).

For the speakers with cochlear implants, repeated-measures ANOVAs were conducted with sessions, consonant place, and feedback as conditions, speaker as a category variable, and variability index as the dependent variable. The preimplant data revealed a significant main effect for participant ($F=6.32$, $p<0.001$, $\eta p^2=0.29$) but consonant place (br, dr, gr) was not significant nor was there an interaction between place and speaker. Consistent with the context variability hypothesis, all of the hearing-impaired participants had greater preimplant variability indices (mean=1.43, sd =0.43) than the normal-hearing controls (mean 1.22, as reported above) with feedback on, although one (FJ) had a variability index distribution similar to the controls (one-sample t -test $p>0.05$). Four of the implant users' preimplant variability index values approached the variability index mean of the feedback-off performance for normal-hearing controls (1.33, as reported above) while three other implant users yielded variability index values higher than those found in either normal hearing (NH) condition. (One implant user (MM) could not be included in the analysis because preimplant data were not collected).

TABLE II. Effect sizes (ηp^2) from the three ANOVAs.

ANOVA effect	NH	CI month	CI year
Speaker	0.43	0.26	0.44
Place	0.13	0.06	0.06
Place X speaker	0.09	0.14	0.13
Feedback	0.12	0.07	0.01
Feedback X speaker	0.43	0.11	0.21

A summary of speaker, place, and feedback effect sizes (partial eta squared, ηp^2) from the ANOVAs for controls and implant users (at two time intervals) is found in Table II. There are substantial differences between the measures for implant users and those for speakers with normal hearing. The speaker main effect for variability index analyses was statistically significant in all three ANOVAs (NH, CI month, and CI year) as was the feedback-by-speaker interaction. Feedback and place (consonant context) were less predictive of allophonic variability in the implant users than in the normal-hearing controls, and the interaction of place-by-speaker was larger for the implant participants, suggesting that certain clusters were more difficult for some implant participants to produce consistently. To summarize, variability indices were greater than 1.0 for all speakers; that is, /r/ was more variable than /a/ under contextual variation. The context variability hypothesis was supported in that varying phonetic context (wabrav, wadrav, wagrav, and warav) yielded more variable acoustics for /r/ from implant users than controls.

2. Longitudinal results

The changes in implant-users' variability indices as they gained experience with prosthetic hearing were analyzed to test the allophonic consistency hypothesis, which postulated that /r/ variability would become more similar to that of controls (i.e., would diminish) over time with the use of the implant. All of the implant participants decreased their variability across sessions so that in the one-month session, one-year session, or both, it was more similar to the results from participants with normal-hearing listening to masking noise in the feedback-off condition. There was a significant main effect of session for implant users during the feedback-off condition ($F=9.9$, $p<0.001$, $\eta p^2=0.07$). There was surprisingly no significant main effect of session during the feedback-on condition, a result that will be explored further below.

a. Individual speaker differences. Despite the indications that prosthetic hearing reduced allophonic variability of /r/ for the group, differences among speakers complicate the picture. The ANOVA results reveal considerable individual variability in the performance of cochlear-implant users across sessions in general and in their reduction in context-induced variability with prosthetic hearing. There was a significant speaker main effect for variability index (feedback on: $F=11.06$, $p<0.001$, $\eta p^2=0.45$ and feedback off: $F=9.64$, $p<0.001$, $\eta p^2=0.40$) and an interaction between speaker and session (feedback on: $F=3.23$, $p<0.001$, $\eta p^2=0.19$ and feedback off: $F=3.35$, $p<0.001$, $\eta p^2=0.19$). The

consonant place by speaker interaction was statistically significant for both the feedback-on ($F=2.035$, $p=0.024$, $\eta^2=0.13$) and feedback-off ($F=2.42$, $p=0.006$, $\eta^2=0.14$) analyses.

Many of the implant users produced some tokens in which the context-induced variability of /r/ was much greater than the variability of /a/ (resulting in variability indices ≥ 2.0) even in the one-year session. Thus, these implant users had variability index values that greatly exceeded those for normal-hearing controls when the latter had access to auditory feedback. A few implant users, however, had some variability index values in postimplant sessions that were lower than the normal-hearing controls' mean. The session by speaker interaction was significant ($\eta^2=0.19$) for both feedback conditions, indicating that some speakers showed a consistent longitudinal trend toward reduced /r/ allophonic variability but others did not.

b. Feedback effects. Overall, three implant users yielded relatively small variability index values and near-normal performance in both feedback conditions by the end of the study. Four implant users showed decreased variability indices in at least one feedback condition as the study progressed. Compared to the preimplant measures, postimplant variability indices showed improvement even when the implant was turned off (feedback-off) as evidenced by a significant session effect ($F=6.91$, $p=0.001$, $\eta^2=0.07$). The main effect for speakers was more pronounced with auditory feedback available from the implant (feedback-on) condition, and the corresponding session effect was not significant ($F=0.62$, $p=0.539$). When the distribution of differences across repetitions of the matched pairs was examined, there were two implant participants in the feedback-on condition who had distributions similar to those obtained in the normal-hearing (feedback-on) condition.

Two observations lead to the inference that the implant users developed stable articulatory targets that were not degraded when their implants were turned off. First, there was a significant main effect of session for implant users during the off condition ($F=9.9$, $p=0.001$, $\eta^2=0.07$). Hence, the variability index improved over the three sessions even in the absence of auditory feedback. Second, the feedback main effect was statistically significant only in the one-month session ($F=7.3$, $p<0.01$) and even there did not account for much variance ($\eta^2=0.07$). Thus, interrupting feedback did not appreciably impair performance. Continued improvements in the variability index in postimplant sessions, therefore, were mostly independent of implant feedback state during testing. Conclusions about group longitudinal trends in variability must be tempered because of the various interaction terms involving participants. In particular, there was no significant effect of session during the feedback-on condition ($\eta^2=0.01$) due to a heightened speaker main effect. In general, however, a decrease in variability index toward normal values was observed for many of the cochlear-implant users, supporting the allophonic consistency hypothesis.

3. Contrasting stop+r with stop alone

To assess changes in the contrast between a singleton stop and that stop followed by /r/ (stop+r/) and test the

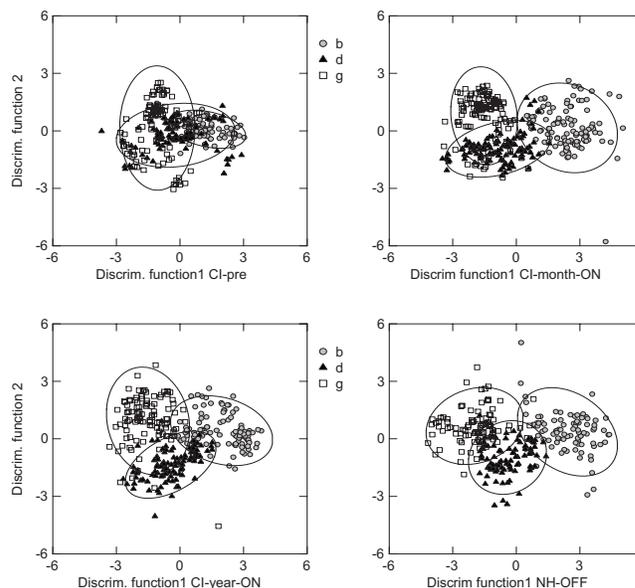


FIG. 5. Results of discriminant analyses of speakers' productions of /b/, /d/, and /g/.

consonant contrast hypothesis (i.e., contrast will be greater for normal-hearing control speakers than for implant users), the distinctiveness of /b/, /d/, and /g/ in wabav, wadav, and wagav was measured using the magnitude of the F3 transition ($\Delta F3$) and the F3–F2 ($kF3-F2$) difference to classify tokens of the phonemes produced.

a. Classification analysis with normal-hearing controls.

The lower right-hand panel of Fig. 5 shows the discriminant analysis classification (90% correct) with the three stop phonemes for the normal-hearing controls in the feedback-off (95 dB SPL masking noise) condition. The less-than-perfect classification was not attributable to a lack of intelligibility in the stop productions by the normal-hearing controls. Rather, it occurred because F2 and F3 measures did not capture all of the crucial information in this contrast. Although adding information about F1 transitions would greatly improve the discriminant analysis results by increasing the separation among /b/, /d/, and /g/, the current analysis focused on F2 and F3 so that the singleton stops could be compared with stop+r/ clusters.

b. Classification analysis with implant users. The upper left-hand panel of Fig. 5 shows the discriminant-analysis classification of the stops produced by the postlingually deaf speakers prior to experience with the cochlear implant. There was considerable overlap among the consonant groups, with 66% of the tokens classified correctly. While /b/ tokens (lightly shaded circles) were tightly clustered and distinct from /g/ (open squares), the /d/ productions (filled triangles) were quite variable. In the one-month data (upper right-hand panel) and one-year data (lower left-hand panel) it can be seen that the result for the discriminant analysis of stop productions by implant users with feedback-on and a year's experience with their implants (89% correct classification) is comparable to that of the normal-hearing controls (90%—lower right-hand panel). The most striking improvement was seen for /d/, which was classified correctly 42% of the time in the preimplant condition but increased to 92% correct at

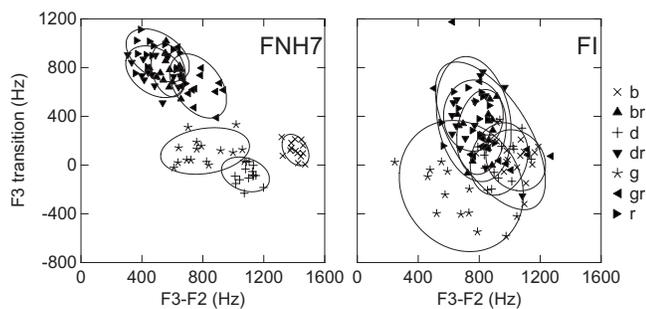


FIG. 6. A comparison of singleton stops (open symbols) vs stop+r tokens (closed, filled symbols) for a normal-hearing speaker with auditory feedback (left panel) and a speaker with hearing loss in the preimplant condition (right panel).

the one-month session and remained relatively distinct from the other two consonants in the one-year analysis.

c. Between-group comparisons and formant measures.

Normal-hearing speakers with better perceptual acuity have been found to produce more distinctive vowel contrasts (Perkell *et al.*, 2004). A listener must perceive certain specific cues to differentiate between a speaker's productions, for example, of wabav and wabrav, so the purpose of these analyses was to evaluate the contrast between the stop productions and their stop+r counterparts under varying feedback conditions. As noted previously, singleton stops can be classified using parameters based on F2 and F3. Figure 6 plots the extent of the F3 transition (see Fig. 1: $\Delta kF3$) against the F3-F2 difference (see Fig. 1: $kF3-F2$) at the beginning of /a/ for a normal-hearing speaker (left-hand panel) with auditory feedback. Tokens for the three plosives (stop+r, filled symbols) overlap with each other (and also with /r/ alone from the warav tokens). When these stops are followed by /r/, they are quite divergent from the singleton stops without /r/. Normal-hearing controls thus reveal a contrast between the stops and the stop+r counterparts with respect to the F3 transition and the difference between F3 and F2 at the consonant/vowel boundary. Before receiving their implants, hearing-impaired speakers generally produced less contrast between stops and their stop+r counterparts; a typical result is seen in the right-hand panel of Fig. 6.

The consonant contrast hypothesis stated that acoustic measures of the difference between stop+r/t/ and singleton stop productions will be greater for control speakers with normal hearing than for implant users. The advantage provided by auditory feedback is expected to yield better phonemic contrast among the stops and between each stop and its stop+r/t/ counterpart. Figure 7 plots the extent of /r/ contrast measured as the difference in F3 transition for singleton stop versus stop+r/t/ productions with the indicated stops. Results are shown for implant users in the one-year session (filled circles) and for normal-hearing controls with their hearing masked (feedback off as clear triangles) or not masked (feedback on as filled triangles). Pooled across the stops, normal-hearing speakers had higher /r/-contrast values (mean=482.1 sd=335.2) than implant users at one year (mean=300.9 sd=339.9) with feedback on (t for independent samples=6.94, $p < 0.001$, $d = 0.53$).

Inspection of Fig. 7 indicates that the normal-hearing

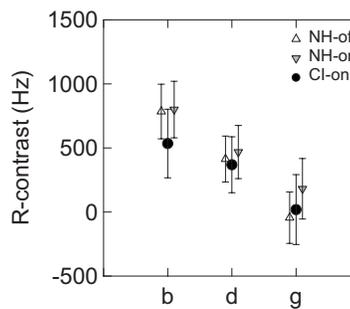


FIG. 7. Extent of /r/ contrast measured as the difference in the F3 transition for singleton stop vs stop+r productions with the indicated stops. Results are shown for implant users in the one-year session (filled circles) and for normal-hearing speakers with their hearing masked (clear triangles) or not masked (filled triangles).

controls, both with and without auditory feedback, as well as the implant users yielded greater contrast for the plosives /b/ and /d/ than for /g/. With the relatively high visibility of the bilabials and the fact that the bilabial+r/t/ blend involves two different primary articulations rather than two overlapping gestures of the tongue (as needed for the velar or alveolar stop+r/t/), deaf participants may have less difficulty contrasting productions of bilabial+r/t/ with bilabial singleton stops.

d. Contrast improvement. Because the preimplant data were obtained with only one feedback condition while the data one month and one year postimplant had two feedback conditions, an ANOVA with session by feedback was not computed. Instead, two repeated-measures ANOVAs were completed. The first had trial variables preimplant, one-month feedback on, and one-year feedback on. The second ANOVA had conditions preimplant, one-month feedback off, and one-year feedback off. There was a significant main effect of session (preimplant, one month, and one yr) on the /r/ versus singleton stop contrast in both the feedback-on ($F = 13.17$, $p < 0.001$, $\eta^2 = 0.14$) and feedback-off ($F = 10.83$, $p < 0.001$, $\eta^2 = 0.11$) conditions. Implant users, therefore, reliably improved contrast over time in both feedback conditions. Implant users varied significantly by session which yielded the largest main effects in these analyses with $\eta^2 = 0.42$ for feedback-on and $\eta^2 = 0.48$ for feedback-off conditions. As illustrated in Fig. 7, when compared to normal-hearing controls exposed to masking noise (feedback off), implant users have poorer /r/ contrast than participants with normal hearing for /b-br/ and /g-gr/ even with auditory feedback from the implant and one year of implant experience. The /d-dr/ contrast is similar for both groups of participants and for both feedback conditions but the contrast was smallest for /g-gr/ for all of the participants even speakers with normal hearing and optimal access to auditory feedback cues.

The speakers with normal hearing produced greater stop+r/t/ versus singleton stop contrast with auditory feedback (Fig. 6, filled triangles, $F = 81.04$, $p < 0.001$, $\eta^2 = 0.45$) than without (unfilled triangles) primarily in the /g/ context. For the implant users, on the contrary, the presence versus absence of feedback did not yield a significant main effect during the one-month ($\eta^2 = 0.009$) or one-year sessions (η^2

=0.004), although it did interact significantly with the speaker variable ($\eta p^2=0.11$ and 0.29 , respectively).

Thus, the contrast between stop+/r/ and stop alone was greater for control speakers with normal hearing than for implant users even when the latter had a year of experience with their implant, although they did improve over the three sessions. Blocking feedback reduced the normal-hearing controls' stop+/r/ contrasts but not those of the implant users.

IV. DISCUSSION

The purpose of the current study was to examine context-dependent allophonic variation in the production of /r/, how it changed longitudinally in cochlear-implant users and how the implant users' production variability compared with that of normal-hearing speakers. This work was intended to address two distinct issues. The first issue is concerned with the nature of motor programming goals in speech production. The second is concerned with the learning that underlies longitudinal reduction in potentially undesirable allophonic variation while maintaining crucial phonemic contrasts. These two issues are joined by a theoretical view of the role of auditory feedback in speech production and how individuals use feedback from a cochlear implant.

A. Goals in a model of speech production

It has been suggested that speakers attempt to satisfy listener demands for clarity but are constrained by economy of effort (Lindblom, 1990; Perkell *et al.* 1997, 2004). As was noted in the descriptions of data from normal-hearing controls, speakers may exhibit a trade-off between clarity and economy of effort and thereby vary widely in the expression of phonetic contrast in their productions. Since /r/ is produced with much greater articulatory variability than almost any other phoneme in English, including the vowel /a/, it was expected and confirmed that the context-induced acoustic variability for /r/ productions would be larger than for /a/ productions for all speakers. This outcome furthered the testing of the context variability hypothesis, which states that varying phonetic context (as in the nonsense words wabrav, wadrav, wagrav, and warav) will yield more variable /r/ productions by implant users than by controls. In their preimplant recording sessions, all of the hearing-impaired participants had greater variability indices than the normal-hearing controls with access to auditory feedback; this finding supports the context variability hypothesis.

Because of the multiple articulatory configurations speakers use in producing /r/ (see Delattre and Freeman, 1968; Guenther *et al.*, 1999; Westbury *et al.*, 1998) and because of relatively crude and frequency-shifted acoustic cues provided by a cochlear prosthesis, implant users may have difficulty in modifying their phoneme-to-articulation maps to produce /r/ appropriately by using their prosthetic hearing. Consequently, implant users would have more difficulty producing semivowels correctly than phonemes with more robust orosensory correlates such as fricative and stop consonants. The consonant contrast hypothesis, which was confirmed, stated that acoustic measures of the difference

between stop+/r/ and singleton stop productions will be greater for speakers with normal hearing than for implant users. Greater contrast for the stop+/r/ blends versus the singleton stops (e.g., b versus br, d versus dr, g versus gr) was found for normal-hearing controls with auditory feedback than for implant users with feedback in all sessions.

B. Longitudinal improvements with implant use

The allophonic consistency hypothesis stated that implant users' acoustic measures of /r/ production after experience with their implants will show a reduction in variability and thereby be more similar to those of controls. Indeed, implant users yielded decreasing values of the variability index as their experience with their implants increased. With just a month's experience with their implants, two of the speakers yielded variability index values comparable to those of the normal-hearing speakers. With a year of prosthetic hearing, seven implant users did so.

Perkell *et al.* (2001) discuss the improved production of /r/ by cochlear-implant speakers, noting that acoustic measures of /r/ and /l/ may overlap in preimplant recordings, but the distributions of the two phonemes become distinct and separate after a year of implant use. The participants described in that study were profoundly deaf and did not have any open-set consonant recognition in an auditory-alone task. By contrast, the speakers in the current study had more residual hearing preimplant and their speech did not seem to have deteriorated as much as had been observed in earlier studies. In fact, the productions of /r/ and /l/ preimplant (analyzed with a similar combination of $\Delta F3$ and $kF3-F2$ measures) in the current implant group were similar to the early postimplant findings of the previous generation of participants. Thus, the acoustic analyses for the current group of speakers had to focus on more subtle acoustic cues than ones that would simply differentiate consonants from one another. To this end, relative acoustic variability, separation among stops in $F2-F3$ space, and the contrast between singleton stops and stop+/r/ clusters were investigated.

The overall aim of a speaker, with or without a hearing loss, is to present the listener with intelligible speech. The information provided by a cochlear implant can help to refine the planning of speech movements so that the speaker can achieve this goal. In the case of the semivowel /r/, if postlingually deafened adults are not able to produce consistently clear productions across phonetic contexts, they are able to improve clarity once they begin receiving acoustic information from a cochlear implant about their own productions and those of others. In terms of the present theoretical framework, profound postlingual hearing loss likely leads to some gradual degradation of sound contrasts and mappings between articulatory and auditory parameters. In the case of American English /r/, such degradation would cause increased variability of $F3$ and increased separation of mean values of $F3$ produced by bunched versus retroflexed articulations. When hearing is partially restored with a cochlear implant, the new unfamiliar auditory stimulation could lead

to some increased variability; however, as mappings and auditory goal regions become refined with time, overall variability and separation of the two allophones decrease.

V. SUMMARY AND CONCLUSIONS

Varying phonetic context (as in the nonsense words, wabrav, wadrav, wagrav, and warav) induced more variability in /r/ productions among implant users than among normal-hearing speakers (supporting the context variability hypothesis). When postlingually deaf speakers were provided some auditory feedback with a cochlear implant, their /r/ productions became more similar to those by hearing controls. Seven out of the eight implant users showed reductions in an index of acoustic variability for /r/ to within normal limits from a preimplant recording to one made one-year postimplant (supporting the allophonic consistency hypothesis). The contrast between stop+r and stop-alone productions was found to be generally larger in speakers with normal hearing than in implant users (supporting the consonant contrast hypothesis). Although implant users showed improvements in the perception and production of stop consonants as well as stop+r blends, their contrast between /b/ and /br/ and between /g/ and /gr/ was still reduced compared to controls, even at one-year postimplant with auditory feedback available.

ACKNOWLEDGMENTS

This research was supported by Grant No. R01-DC003007 from the National Institute on Deafness and Other Communication Disorders, National Institutes of Health. We are grateful to Dr. Donald Eddington of the Massachusetts Eye and Ear Infirmary and Dr. Daniel Lee of the Department of Otolaryngology-Head and Neck Surgery, University of Massachusetts Medical Center for referring cochlear-implant users to our study. We also wish to thank Advanced Bionics and Cochlear for donating research devices for our use.

¹Partial eta squared, η^2 , was used as an effect size measure to assess the strength of the association between the sample main effect and the dependent variable. It is calculated as the proportion of the sums of squares due to the effect divided by the sums of squares due to the effect plus the sums of squares due to error (Tabachnik and Fidell, 1996). As partial eta squared approaches 1.0, it reflects more variance accounted for by the effect under study.

- Adler-Bock, M., Bernhardt, B. M., and Bacsfalvi, P. (2007). "The use of ultrasound in remediation of North American English /r/ in two adolescents," *Am. J. Speech Lang. Pathol.* **16**, 128–139.
- Blamey, P. J., Dowell, R. C., Brown, A. M., Clark, G. M., and Seligman, P. M. (1987). "Vowel and consonant recognition of cochlear implant patients using formant-estimating speech processors," *J. Acoust. Soc. Am.* **82**, 48–57.
- Boyce, S. and Espy-Wilson, C. Y. (1997). "Coarticulatory stability in American English /r/," *J. Acoust. Soc. Am.* **101**, 3741–3753.
- Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., and Tohkura, Y. (1999). "Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech perception," *J. Acoust. Soc. Am.* **101**, 2299–2310.
- Delattre, P., and Freeman, D. (1968). "A dialect study of American R's by x-ray motion picture," *Language* **44**, 29–68.
- Espy-Wilson, C. Y. (1992). "Acoustic measures for linguistic features distinguishing the semivowels /w j r l/ in American English," *J. Acoust. Soc. Am.* **92**, 736–757.

- Guenther, F. H. (1995). "Speech sound acquisition, coarticulation and rate effects in a neural network model of speech production," *Psychol. Rev.* **102**, 594–621.
- Guenther, F. H., Espy-Wilson, C. Y., Boyce, S. E., Matthies, M. L., Zandipour, M., and Perkell, J. (1999). "Articulatory tradeoffs reduce acoustic variability during American English /r/ production," *J. Acoust. Soc. Am.* **105**, 2854–2865.
- Guenther, F. H., Ghosh, S. S., and Tourville, J. A. (2006). "Neural modeling and imaging of the cortical interactions underlying syllable production," *Brain Lang* **96**, 280–301.
- Guenther, F. H., Hampson, M., and Johnson, D. (1998). "A theoretical investigation of reference frames for the planning of speech movements," *Psychol. Rev.* **105**, 611–633.
- Ingram, J. C. L., and Park, S.-G. (1998). "Language, context, and speaker effects in the identification and discrimination of English /r/ and /l/ by Japanese and Korean listeners," *J. Acoust. Soc. Am.* **103**, 1161–1174.
- Kishon-Rabin, L., Taitelbaum, R., Tobin, Y., and Hildesheimer, M. (1999). "The effects of partially restored hearing on the speech production of postlingually deafened adults with multichannel cochlear implants," *J. Acoust. Soc. Am.* **106**, 2843–2857.
- Lane, H., Matthies, M., Denny, M., Guenther, F., Perkell, J., Stockmann, E., Tiede, M., Vick, J., and Zandipour, M. (2007). "Effects of short- and long-term changes in auditory feedback on vowel and sibilant contrasts," *J. Speech Lang. Hear. Res.* **50**, 913–927.
- Lane, H., Matthies, M., Perkell, J., Vick, J., and Zandipour, M. (2001). "The effects of changes in hearing status in cochlear implant users on the acoustic vowel space and coarticulation," *J. Speech Lang. Hear. Res.* **44**, 552–563.
- Lane, H., Wozniak, J., Matthies, M. L., Svirsky, M. A., and Perkell, J. S. (1995). "Phonemic resetting versus postural adjustments in the speech of cochlear implant users: An exploration of voice-onset time," *J. Acoust. Soc. Am.* **98**, 3096–3106.
- Lindblom, B. (1990). "Explaining phonetic variation: A sketch of the H and H is theory," in *Speech Production and Speech Modeling*, edited by W. J. Hardcastle and A. Marchal (Kluwer, Dordrecht), pp. 403–439.
- Matthies, M. L., Svirsky, M. A., Lane, H., and Perkell, J. S. (1994). "A preliminary study of the effects of cochlear implants on the production of sibilants," *J. Acoust. Soc. Am.* **96**, 1367–1373.
- Matthies, M. L., Svirsky, M. A., Perkell, J. S., and Lane, H. (1996). "Acoustic and articulatory measures of sibilant production with and without auditory feedback from a cochlear implant," *J. Speech Hear. Res.* **39**, 936–946.
- McGowan, R. S., Nittrouer, S., and Manning, C. J. (2004). "Development of [ɹ] in young, Midwestern, American children," *J. Acoust. Soc. Am.* **115**, 871–884.
- McKay, C. M., and McDermott, H. J. (1993). "Perceptual performance of subjects with cochlear implants using the spectral maxima sound processor (SMSP) and the mini speech processor (MSP)," *Ear Hear.* **14**, 350–367.
- Nieto-Castanon, A., Guenther, F. G., Perkell, J. S., and Curtin, H. D. (2005). "A modeling investigation of articulatory variability and acoustic stability during American English /r/ production," *J. Acoust. Soc. Am.* **117**, 3196–3212.
- Ong, D., and Stone, M. (1998). "Three-dimensional vocal tract shapes in /r/ and /l/: A study of MRI, ultrasound, electropalatography, and acoustics," *Phonoscope* **1**, 1–13.
- Perkell, J. S., Guenther, F. H., Lane, H., Matthies, M., Perrier, P., Vick, J., and Zandipour, M. (2001). "Planning and auditory feedback in speech production," *Speech Motor Control in Normal and Disordered Speech: Proceedings of the 4th International Speech Motor Conference (Nijmegen: The Netherlands)*, pp. 5–11.
- Perkell, J. S., Guenther, F. H., Lane, H., Matthies, M. L., Stockmann, E., Tiede, M., and Zandipour, M. (2004). "The distinctness of speakers' productions of vowel contrasts is related to their discrimination of the contrasts," *J. Acoust. Soc. Am.* **116**, 2338–2344.
- Perkell, J. S., Lane, H., Svirsky, M. A., 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. L., Lane, H., Guenther, F., Wilhelms-Tricarico, R., Wozniak, J. F., and Guiod, P. (1997). "Speech motor control: Acoustic goals saturation effects, auditory feedback, and internal models," *Speech Commun.* **22**, 227–250.
- Stevens, K. N. (1998). *Acoustic Phonetics* (MIT Press, Cambridge, MA).

- Tabachnik, B. G., and Fidell, L. S. (1996) *Using Multivariate Statistics* 3rd ed. (HarperCollins College Publishers, New York).
- Waldstein, R. (1990). "Effects of postlingual deafness on speech production: Implications for the role of auditory feedback," *J. Acoust. Soc. Am.* **88**, 2099–2114.
- Westbury, J. R., Hashi, M., and Lindstrom, M. (1998). "Differences among speakers in lingual articulation for American English /r/," *Speech Commun.* **26**, 203–226.
- Wilson, B., Lawson, D., Zerbi, M., Finley, C., and Wolford, R. (1995). "New processing strategies in cochlear implantation," *Am. J. Otolaryngol.* **16**, 669–681.
- Zhou, X., Espy-Wilson, C. Y., Boyce, S., Tiede, M., Holland, C., and Choe, A. (2008). "A magnetic resonance imaging-based articulatory and acoustic study of "retroflex" and "bunched" American English /r/," *J. Acoust. Soc. Am.* **123**, 4466–4481.