6-14-2000

Perceptual Calibration of F0 Production: Evidence from Feedback Perturbation

Jeffery A. Jones
Wilfrid Laurier University, jjones@wlu.ca

Kevin Munhall
Queen's University - Kingston, Ontario

Recommended Citation

Jones, J.A. & Munhall, K.G. (2000). Perceptual calibration of F0 production: Evidence from feedback perturbation. The Journal of the Acoustical Society of America 108, 1246. doi: 10.1121/1.1288414

Perceptual calibration of *F*0 production: Evidence from feedback perturbation

Jeffery A. Jones
Department of Psychology, Queen's University, Kingston, Ontario K7L 3N6, Canada

K. G. Munhall
Department of Psychology and Department of Otolaryngology, Queen's University, Kingston, Ontario K7L 3N6, Canada

(Received 3 March 2000; accepted for publication 14 June 2000)

Hearing one's own speech is important for language learning and maintenance of accurate articulation. For example, people with postlinguistically acquired deafness often show a gradual deterioration of many aspects of speech production. In this manuscript, data are presented that address the role played by acoustic feedback in the control of voice fundamental frequency (*F*0). Eighteen subjects produced vowels under a control (normal *F*0 feedback) and two experimental conditions: *F*0 shifted up and *F*0 shifted down. In each experimental condition subjects produced vowels during a training period in which their *F*0 was slowly shifted without their awareness. Following this exposure to transformed *F*0, their acoustic feedback was returned to normal. Two effects were observed. Subjects compensated for the change in *F*0 and showed negative aftereffects. When *F*0 feedback was returned to normal, the subjects modified their produced *F*0 in the opposite direction to the shift. The results suggest that fundamental frequency is controlled using auditory feedback and with reference to an internal pitch representation. This is consistent with current work on internal models of speech motor control. © *2000 Acoustical Society of America.*
[S0001-4966(00)03009-5]

PACS numbers: 43.70.Aj, 43.66.Hg [AL]

I. INTRODUCTION

There are many indications that fluent speech is controlled through the use of sophisticated internal representations as well as feedback processed on-line. For example, people with postlinguistically acquired deafness often show a deterioration of many aspects of speech production. Problems related to intensity and pitch control, as well as intonation, stress, and rate of speech are commonly seen quite soon after hearing loss. However, only after longer periods of deafness will variability in the production of vowels and consonants be observed (Cowie and Douglas-Cowie, 1992).

The finding that the precision of vowel and consonant production persists unaltered for a relatively long time after deafness onset supports the existence of a well-formed neural mapping between the motor system and the acoustic signals for segments. On the other hand, the finding that deafness more rapidly affects production parameters such as pitch and intensity implies that the mechanisms involved in suprasegmental control may be different than those for the control of segment production (Perkell *et al.*, 1997). The control of these parameters may be more directly sensitive to acoustic feedback. For example, speakers exposed to loud noise spontaneously and immediately compensate by increasing the volume of their speech (Lane and Tranel, 1971).

Uncovering how such a complex control system operates and determining the relative roles played by feedback and central representations is a daunting task requiring both empirical and modeling work. Recently, there has been considerable interest in the role of feedback and ''internal models'' in motor control in general. Internal models are hypothesized neural representations of the spatial (kinematic), force (dynamic), and/or proprioceptive characteristics of movements that could be used by the nervous system to predict movement outcome. These predictive models could provide internal feedback to planning and control systems without the delays associated with natural proprioceptive feedback [see Miall and Wolpert (1996); Kawato (1999) for discussions of the many roles of internal models in movement control].

Evidence for the existence of internal models comes primarily from the study of arm and hand movements. For example, Johanssen and Westling (1984) and Flanagan and Wing (1993) have shown that when grasping an object with the hand, grip force changes in synchrony with changes in load forces on the object. This synchrony could only result from control that predicts the loads on the object and thus the grip force needed to hold the object.

Although internal models can potentially reduce the need for closed-loop control, feedback still plays an important role in their acquisition and maintenance. Subjects exposed to novel conditions can acquire new internal models. Several investigators have shown that subjects exposed to an artificial force field while making point-to-point movements adapt and eventually make arm movements with natural trajectories. For example, Shadmehr and Mussa-Ivaldi (1994) had subjects move a robot manipulandum to targets while the robot imposed forces. Initially, the trajectories produced by the subjects were distorted; however, with practice, the subjects produced movement paths quite similar to movements produced prior to exposure to the force field. When the forces were suddenly removed, the subjects showed aftereffects in their movements for a few trials. For these trials, subjects moved as if they were encountering the experimental force field even though it was no longer present. This pattern of behavior suggests that subjects relearned the mapping between the kinematics of arm movements and the forces needed to control trajectories. In other words, they constructed a new internal model to accomplish the reaching task under the novel force conditions.

The aftereffects shown by Shadmehr and Mussa-Ivaldi are a form of sensorimotor adaptation similar to that observed with visual and vestibular perturbation paradigms [see Welch (1986) for a review]. For example, Held (1965) showed that subjects wearing prisms that displace the visual field quickly relearned the mapping between the visual space and the motor system. Initially, the subjects made reaching errors in the direction of the prism displacement. After a number of practice trials they returned to normal accuracy and normal movement speed. However, when the prisms were removed they made reaching errors in the opposite direction to the prism displacement. These data have been interpreted as evidence for a learned mapping between the movement and perceptual systems.

Similar experiments have been conducted in speech production research addressing the learned mappings between vocal tract movement and the resulting acoustics. The considerable variability that exists in vocal tract morphology means that talkers must learn the unique acoustic characteristics of their own vocal tract in order to produce the sounds of their language. In formal models of acoustic-articulatory mappings (Guenther, 1994; Hirayama *et al.*, 1994; Jordan, 1990, 1996; Kawato *et al.*, 1987) acoustic feedback plays a number of possible roles: (1) For speech sound development in children and adults and for learning new vocal tract arrangements, acoustic feedback provides the primary information about target achievement and thus is the vehicle for learning. (2) For fine motor control, the sound of the speaker's voice is used in closed-loop control of articulation. (3) For motor planning and control, the vocal acoustics provides an ongoing calibration of internal models of the speech motor system.

In this paper we explore the relative importance of the third role: the use of acoustic feedback in calibrating an internal model for the control of speaking fundamental frequency ($F0$). $F0$ is determined partly by individual anatomy and physiology and partly by a control system that relies on feedback to achieve a pitch ''target'' (Titze, 1994). The biomechanical and physiological contributions to the fundamental frequency of vocal fold vibration include the mass of the folds, the subglottic lung pressure, and tension on the folds from a network of muscles such as the cricothyroid and vocalis muscles. These biophysical factors are controlled by a complex network of cortical and brainstem centers (Larson, 1988) as well as proprioceptive (Kirchner and Wyke, 1965) and auditory (Sapir *et al.*, 1983) reflex mechanisms.

During normal conversation, the pitch of the voice varies as a function of speaking intensity, prosodic pattern, emotionality and speaking rate, but for any given individual this frequency range varies around an ''habitual'' vocal pitch (Zemlin, 1981). In this paper we test the extent to which this habitual pitch is controlled by an internal $F0$ target. As others have previously, we use a modified feedback approach. Several studies have demonstrated that when subjects hear their $F0$ feedback suddenly raised or lowered artificially, they compensate by shifting their pitch in the opposite direction (e.g., Burnett *et al.*, 1998; Kawahara, 1995). In the protocol used in this experiment, vocal pitch feedback was slowly shifted up or down in frequency without subjects' awareness. Our primary aim was to demonstrate adaptation to modified pitch feedback following return to normal feedback conditions.

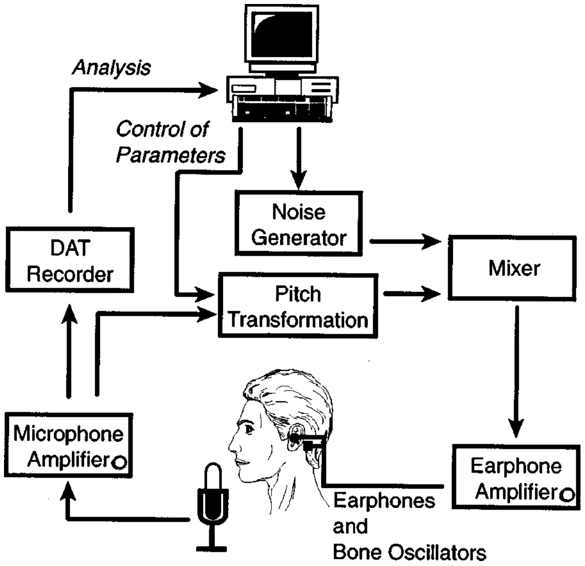

FIG. 1. Schematic of experimental acoustic feedback setup.

There are relatively few reports of auditory adaptation in speech production research. Houde and Jordan (1998) gave speakers real-time auditory feedback in which the formants they were producing were shifted enough to change the vowel's phonetic identity. Over many trials, Houde and Jordan found that speakers modified their vowel productions to compensate for the ongoing feedback transformations. In addition, the modified productions persisted in the absence of feedback, indicating an adaptive response involved in updating the acoustic-motor representation. In our experiment, if talkers show aftereffects of modified feedback conditions this will be clear evidence that habitual speaking $F0$ is controlled relative to an internally represented reference frequency.

II. METHOD

A. Subjects

Eighteen male speakers of Canadian English between 18 and 30 years of age (mean of 22.4 years) participated. The participants reported no hearing, speech, or language problems.

B. Apparatus

Figure 1 depicts the experimental setup. Utterances were recorded using a Telex PH-20 microphone. Prior to pitch shifting, the signals were amplified (Tucker-Davis MA2 microphone amplifier) and filtered (Tucker-Davis FT6-2) with a 9-kHz frequency cutoff. An Eventide Ultra Harmonizer (H3000-D/SX) transformed the pitch of the signals. The pitch-shift processing introduced only a small delay (3–4 ms). Trial initiation and the pitch processor were controlled by a computer. To reduce the amount of natural acoustic feedback, the pitch-shifted signals were mixed with pink noise (Grason-Stadler 901B) and multi-speaker babble (Auditec, St. Louis) and then attenuated by a Yorkville reference amplifier (model SR 300). The level of the masking noise was 75 dB SPL. Subjects received the auditory feedback through Etymotic (ER-2) earphones and through Radioear Model B-71 bone oscillators positioned on the left and right mastoid processes. Both the altered and unaltered signals were recorded at 48 kHz on DAT.

The experimental sessions recorded on DAT were later low-pass filtered (with a 5-kHz cutoff) and digitized with a sampling rate of 11 kHz. The fundamental frequency of utterances during each trial was calculated using an algorithm incorporated in the commercial software package CspeechSP (Milenkovic and Read, 1995). The median frequency value for each utterance made during the 3-s interval was used for subsequent analyses.

C. Procedure

Each subject was seated in a small room in front of a computer monitor. Bone oscillators were fixed to the mastoid processes using a flexible headband. The subject also wore a helmet that held a microphone at a fixed distance (7 cm) from the mouth. The transducers for the earphones were attached to a velcro strap around the subject's neck and foam inserts were positioned comfortably in the subject's ear canals.

Depicted on a computer monitor in front of the subject was the word ''awe.'' Below the word was a countdown from 3 to 0 s. Subjects were asked to produce the vowel /ɑ/ (represented orthographically by the word ''awe'') for the duration of the countdown and then to click on an icon at the bottom of the screen with a mouse to initiate the next trial.

The subjects were asked to try to produce the vowel the same way from trial to trial. However, the experimenter made no references to pitch or other voice characteristics and subjects were not made aware of the nature of the experimental manipulation.

D. Experimental design

There were three conditions in the experiment: a ''shift-up,'' ''shift-down,'' and ''control'' condition. Subjects participated in all three feedback conditions and the order of conditions was counterbalanced across subjects. The experimental sessions took place on different days to avoid vocal fatigue.

In the shift-up condition, subjects first produced ten utterances while receiving normal feedback. These ten utterances were later used to establish the subject's baseline $F0$ for the session. Following the ten baseline trials, subjects produced another 100 utterances. For each successive utterance, the pitch of their auditory feedback was increased by one cent. These trials were followed by 20 trials in which feedback was maintained at 100 cents above the subjects' true $F0$. Finally, subjects performed ten trials in which the feedback they heard was normal, that is, unaltered. From the subject's viewpoint, 140 trials were recorded without interruption. The stages of the experiment were implemented without any formal indication of changes in the feedback conditions.

The shift-down condition was conducted in exactly the same manner as in the up condition except subjects were exposed to decreasing pitch feedback to a maximum of -100 cents after the initial ten baseline trials. The pitch of their auditory feedback was decreased by one cent on each of 100 trials. Subjects were then exposed to 20 trials in which feedback was maintained at 100 cents below their true $F0$. These were followed by ten trials in which normal feedback was given.

In addition to the experimental manipulations, subjects also participated in a control condition in which they produced an equivalent number of trials without frequency manipulations. The condition was an attempt to control for $F0$ changes that may result from repeatedly producing the same sound in an experimental setting. Because insert earphones were used, the ''normal feedback'' condition was in reality a small transformation of the normal auditory feedback. Pinnae reflections change the quality of normal airborne feedback but were not present in our auditory feedback through the earphones.

The median frequency value for each utterance made during each trial was obtained and converted to cents based on the following formula:

$$\text{Cents}=100(12\log_2(F/B)),$$

where F is the median frequency for the utterance during the trial, B is the average of the median frequencies for the ten utterances during the baseline phase at the beginning of the experimental session.

III. RESULTS

Figure 2(a) presents the data in Hertz averaged across subjects for the control, shift-up, and shift-down conditions. As can be seen, when the pitch feedback is shifted down, the subjects raised their pitch compared to when the pitch feedback is shifted up. When the feedback was returned to normal, the $F0$ in both shift conditions changed. The mean differences between the final 20 training trials and the final 10 normal feedback trials of the experiment for the control, shift-up, and shift-down conditions were quite small in absolute terms (0.35, -2.8, and 1.5 Hz, respectively) but reliable. In the shift-up condition, $F0$ increased while the $F0$ in the shift-down condition dropped, generating a significant interaction $[F(2,34)=8.1,\ p=0.001]$. After normal feedback was returned for the test trials, the mean pitch for the shift-up condition increased significantly $(p=0.002)$ and shift-down conditions decreased significantly $(p=0.047)$.

Since the experimental sessions for each subject took place on different days, and because the subjects could have different baseline $F0$'s during the different sessions, we also converted the data to cents for comparison within and be-

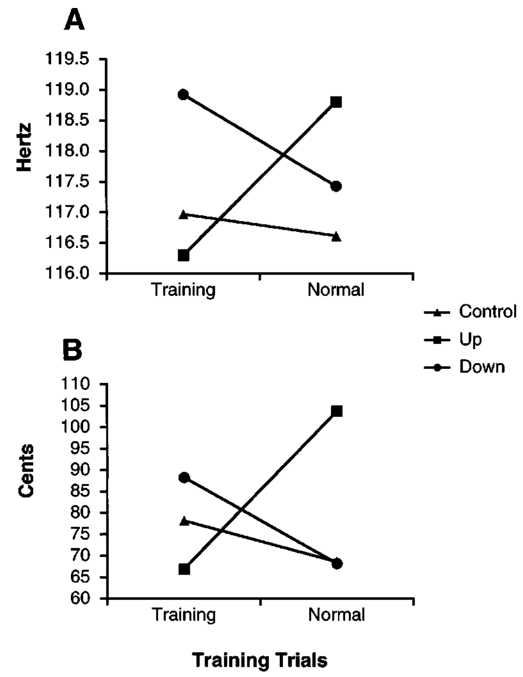

FIG. 2. Average fundamental frequency in Hertz (a) and cents (b) for the three feedback conditions (normal control, shifted up, shifted down) during the final 20 trials of the training period and when feedback was returned to normal.

tween subjects. The conversion of frequency values to cents served to normalize the data with respect to $F0$ baseline trials produced at the beginning of each experimental session.

Figure 2(b) shows the mean $F0$ in cents for the last 20 trials of the training period (i.e., $F0$ shifted 100 cents) and the $F0$ for the final 10 trials of the experiment (i.e., with normal $F0$ feedback). As can be seen, the data in cents show the same pattern. In the shift-up condition, $F0$ increased in response to normal feedback while the $F0$ in the shift-down condition dropped, generating a significant interaction $[F(2,34)=6.2,\ p=0.005]$.

Figure 3 shows the pitch patterns in cents for the three conditions during the 120 training trials. As can be seen, all conditions show an increase in $F0$ with respect to their baseline values with all conditions showing a significant linear trend $[F(1,17)=9.63,\ p=0.006]$. However, the shape of the function for each condition differed. The pitch values for the shift-down condition diverge from the control condition and become higher during the training than the other conditions. Conversely, the pitches during the shift-up condition diverged downward from the control condition. The pitches produced in the three conditions reached their maximum separation at the end of the 100 trials of training. This separation is reduced slightly and the pitch values for the shift-up and shift-down conditions tend to converge toward the control condition performance during the final 20 training trials in which pitch feedback was maintained at 100 cents above or below subjects' true $F0$. Figure 4 shows the produced and heard $F0$ in Hertz for the same 120 training trials. As noted above, the $F0$ shifts were intentionally small to avoid detection by subjects.

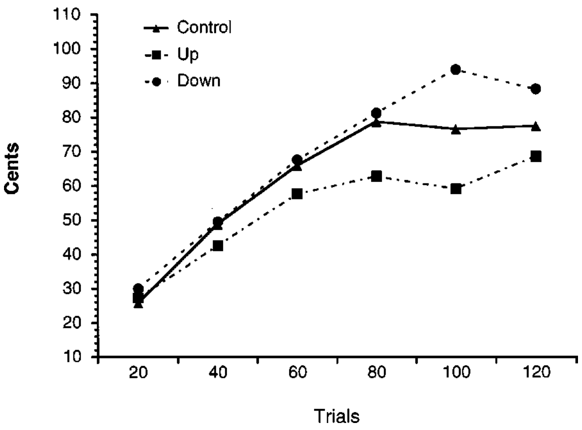

FIG. 3. Average fundamental frequency in cents as a function of blocks of 20 trials during the training period for the three feedback conditions (normal feedback control, shifted up, shifted down).

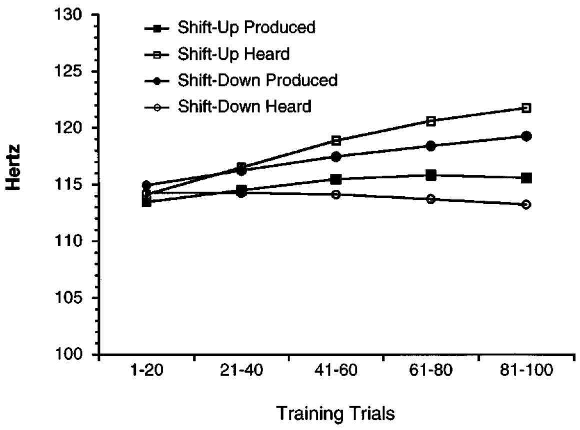

FIG. 4. Average produced and heard fundamental frequency in Hertz as a function of blocks of 20 trials during the training period for the two altered feedback conditions (shifted up, shifted down).

Subjects produced the vowels under loud auditory feedback conditions and it is possible that the observed $F0$ changes could have been due to shifts in speaking volume over the course of the training period. To test for this, the root-mean-square (rms) amplitude of the vowels was computed for the initial baseline ten trials and the final ten test trials for each of the three feedback conditions. An ANOVA showed no effects of training time [initial baseline versus final test trials: $F(1,17)=0.001,\ p>0.5$], feedback condition [upward, downward, control: $F(2,34)=0.592,\ p>0.5$], nor the interaction of these two variables $[F(2,34)=0.113,\ p>0.5]$. Thus pitch changes associated with speaking volume adjustment do not account for the observed $F0$ modifications.

In post-experimental interviews, none of the subjects reported being aware of the gradual shifts in $F0$ feedback. While all subjects were aware that something had been manipulated when $F0$ was returned to normal at the end of the training conditions, they were at a loss to explain what had transpired.

IV. DISCUSSION

The data in this study show two related effects of feedback transformation on $F0$. During the training period subjects *compensated* for the pitch shifts in an apparent attempt to maintain habitual pitch targets under feedback control. When pitch was shifted up subjects lowered their pitch relative to a control condition; when pitch was shifted down they raised their pitch relative to the control. The subjects also showed evidence of sensorimotor *adaptation*. Aftereffects resulted from the relatively short period of exposure to the altered $F0$ feedback. When subjects heard $F0$ feedback that was higher than their true $F0$, the pitch of their voice increased when they were unexpectedly given normal, unaltered auditory feedback. The opposite effect was observed when subjects heard $F0$ feedback lower than it actually was.

Multiple components must be involved in the vocal pitch control system responsible for this behavior. Some have suggested that there is an ''optimum'' pitch range (Zemlin, 1981) determined by the anatomy and physiology of the vocal mechanism. Since fundamental frequency varies as a function of prosody, speaking volume, social situation, emotional state, etc., a mechanism must exist for the controlled modulation of this ''natural level.'' A number of studies (e.g., Kawahara, 1995; Larson *et al.*, 2000) and the present compensation data have shown that auditory feedback can be used in a closed-loop fashion to control fundamental frequency. Presumably, talkers also use kinesthetic feedback or its perceptual concomitant, vocal effort, to aid in this control. In addition, the present data suggest that some type of internal model or representation plays a role in the long-term calibration of vocal pitch. This conclusion is supported by the aftereffect or adaptation data.

The current adaptation results are analogous to those found in classic prism experiments (e.g., Held, 1965). Following a training period wearing displacing prisms, subjects make errors for a short time in the opposite direction of the prism displacement. This aftereffect is widely considered to result from adaptation. The mechanism of the adaptation remains controversial, but it is clear, that at some level a remapping between retinal space and body space must occur. In our data, the subjects acted as if a remapping between perceived and produced pitch had taken place.

It is difficult to determine at what level this $F0$ remapping is taking place. It may be that a representation of a base or neutral pitch level is modulated during the training phase. However, habitual pitch (i.e., average $F0$ from day to day) appears to be quite variable and thus a narrow $F0$ target range seems unlikely. In a study by Coleman and Markham (1991), habitual pitch was found to vary as much as plus or minus three semitones or approximately 18% (cf. Titze, 1994). On the other hand, subjects may attempt to match pitch with perceived vocal effort or kinesthetic feedback (Guenther, 1994). When pitch is shifted up or down, it may not be the absolute pitch value that drives compensation, but the discrepancy between the kinesthetic and auditory feedback.

In this study, there was a tendency for subjects to gradually increase their pitch during the experimental session independent of the feedback condition (see Fig. 3). Since we did not ask subjects to maintain a particular loudness level one possible explanation for this pattern is that subjects increased their speaking volume during the session, causing an increase in pitch (Gramming *et al.*, 1988). However, there was no significant difference between the rms amplitudes of the utterances during the sessions. It is also possible that the increased pitch is related to vocal fatigue developed over the session. Unfortunately, there is no established method for assessing vocal fatigue from the acoustic record (Titze, 1994). It should be noted that the tendency for vocal pitch to increase even when there is no feedback manipulation underscores the value of the control condition. If only one shift condition had been tested with no control group, the effect size of the feedback manipulation could not be determined.

While this trend for pitch to increase is controlled for, it obscures the sensitivity of subjects to the pitch shifts. None of the subjects was consciously aware that the pitch feedback was being modified but their control systems ultimately responded to the changes. The $F0$ patterns in each training condition could be used to test at what size of pitch shift the $F0$ control system began to compensate. Figures 3 and 4 show that for small shifts at the beginning of the training period all three conditions show a similar function and only with larger shifts do the experimental conditions diverge from the control. For upward shifts, the subjects diverged from the control condition earlier than for the downward shifts. Unfortunately, all conditions were tested on different days, so the conditions producing the gradual increase may not have been constant across days. This question must await further study.

Klatt (1973) has shown that subjects can make quite fine perceptual discriminations (between 0.3 and 0.5 Hz) in the $F0$ of synthesized vowels with flat $F0$ contours. When the vowels were synthesized with linearly decreasing $F0$ or as a diphthong with a natural $F0$ contour, the discrimination threshold rose above 2 Hz. Since the final pitch shifts in the experiment were small and in this range, our subjects may have been operating at their perceptual limen. However, evidence from visual-motor control (e.g., Milner and Goodale, 1995) indicates that there can exist perceptual systems for the control of action that are separate from the perceptual system used in categorical judgements. A recent magnetoencephalography study has provided evidence that suggests that this also could be true for the auditory system in speech. Houde *et al.* (2000) have shown that the auditory cortex responds differently to hearing one's own speech while producing it versus listening to recordings of one's own speech. Thus the Klatt threshold data may not be relevant to the issue of sensitivity to feedback modification.

The kind of short-term learning that was observed in this study has been reported in many speech studies previously. Subjects adapt to various static [e.g., bite block (McFarland and Baum, 1995; palatal prostheses, Baum and McFarland, 1997; Hamlet and Stone, 1976, 1978; Hamlet *et al.*, 1978)] or dynamic (Gracco and Abbs, 1986) physical perturbations and auditory feedback transformations (e.g., Houde and Jordan, 1998). We have chosen to consider this learning in the context of internal models. The general concept of an internal model is not a new one. Similar roles have been played by motor programs (e.g., Keele, 1968), efference copy (e.g., von Holst and Mittelstaedt, 1950) and feedforward control (e.g., Arbib, 1981). In proposals about speech production, internal models for vocal tract geometry, kinesthetic, and acoustic mappings have been postulated. Our work suggests that the acoustic mapping must be differentiated to include an $F0$ model as well [see Kawato (1999) for a discussion of multiple internal models].

This suggestion is consistent with evidence from clinical populations. Post-linguistically deafened individuals often have difficulty producing normal intonations soon after their hearing is lost (Cowie and Douglas-Cowie, 1992). Perkell and his colleagues have shown that better $F0$ control is achieved after activation of cochlear implants (Perkell *et al.*, 1992). Perkell *et al.* (1997) have proposed that the auditory system uses information regarding conditions for intelligibility (e.g., ambient noise, social context) in a closed-loop fashion to rapidly make adjustments in $F0$ and vocal intensity. The data in the present study suggest that in addition to this closed-loop control, auditory feedback may also play a role in establishing a baseline for the controlled parameters.

ACKNOWLEDGMENTS

This work was funded by NIH Grant No. DC-00594 from the National Institute of Deafness and other Communications Disorders and NSERC. Chuck Larson and David McFarland made helpful comments on an earlier draft of this manuscript.

Arbib, M. A. (**1981**). ''Perceptual structures and distributed motor control,'' *Handbook of Physiology—The Nervous System II*, edited by V. B. Brooks (American Physiology Society, Bethesda, MD).

Baum, S. R., and McFarland, D. H. (**1997**). ''The development of speech adaptation to an artificial palate,'' J. Acoust. Soc. Am. **102**, 2353–2359.

Burnett, T. A., Freedland, M. B., Larson, C. R., and Hain, T. C. (**1998**). ''Voice $F0$ responses to manipulations in pitch feedback,'' J. Acoust. Soc. Am. **103**, 3153–3161.

Coleman, R. F., and Markham, I. W. (**1991**). ''Normal variations in habitual pitch,'' J. Voice **5**, 173–177.

Cowie, R., and Douglas-Cowie, E. (**1992**). ''Postlingually acquired deafness,'' in *Trends in Linguistics, Studies and Monographs* (Mouton de Gruyter, New York), Vol. 62.

Flanagan, J. R., and Wing, A. M. (**1993**). ''Modulation of grip force with load force during point-to-point arm movements,'' Exp. Brain Res. **95**, 131–143.

Gracco, V. L., and Abbs, J. H. (**1986**). ''Variant and invariant characteristics of speech movements,'' Exp. Brain Res. **65**, 156–166.

Gramming, P., Sundberg, J., Ternström S., Leanderson, R., and Perkins, W. H. (**1988**). ''Relationship between changes in voice pitch and loudness,'' J. Voice **2**, 118–126.

Guenther, F. (**1994**). ''A neural network model of speech acquisition and motor equivalent production,'' Biol. Cybern. **72**, 43–53.

Hamlet, S., and Stone, M. (**1976**). ''Compensatory vowel characteristics resulting from the presence of different types of experimental dental prosthesis,'' J. Phonetics **4**, 199–218.

Hamlet, S., and Stone, M. (**1978**). ''Compensatory alveolar consonant production induced by wearing a dental prosthesis,'' J. Phonetics **6**, 227–248.

Hamlet, S., Stone, M., and McCarty, T. (**1978**). ''Conditioning dentures viewed from the standpoint of speech adaptation,'' J. Prosthet. Dent. **40**, 60–66.

Held, R. (**1965**). ''Plasticity in sensory-motor systems,'' Sci. Am. **213**, 84–94.

Hirayama, M., Vatikiotis-Bateson, E., and Kawato, M. (**1994**). ''Inverse dynamics of speech motor control,'' in *Advances in Neural Information Processing Systems*, edited by S. J. Hanson, J. D. Cowan, and C. L. Giles (Morgan Kaufmann, San Mateo, CA), Vol. 6, pp. 1043–1050.

Houde, J. F., and Jordan, M. I. (**1998**). ''Sensorimotor adaptation in speech production,'' Science **279**, 1213–1216.

Houde, J., Nagarajan, S., and Mersenich, M. (**2000**). ''Modulation of auditory cortex during speech production: An MEG study,'' *Proceedings of the 5th Seminar on Speech Production: Models and Data*, pp. 249–252.

Johansson, R. S., and Westling, G. (**1984**). ''Roles of glabrous skin receptors and sensorimotor memory in automatic-control of precision grip when lifting rougher or more slippery objects,'' Exp. Brain Res. **56**, 550–564.

Jordan, M. I. (**1990**). ''Motor learning and the degrees of freedom problem,'' *Attention and Performance*, edited by M. Jeannerod (Lawrence Erlbaum, Hillsdale, NJ), Vol. XIII, pp. 796–836.

Jordan, M. I. (**1996**). ''Computational aspects of motor control and motor learning,'' *Handbook of Perception and Action*, edited by H. Heuer and S. Keele (Academic, New York), Vol. 2, pp. 71–120.

Kawahara, H. (**1995**). ''Transformed auditory feedback: The collection of data from 1993.1 to 1994.12 by a new set of analysis procedures,'' ATR Technical Report, TR-H-120.

Kawato, M. (**1999**). ''Internal models for motor control and trajectory planning,'' Curr. Opin. Neurobiol. **9**, 718–727.

Kawato, M., Furukawa, K., and Suzuki, R. (**1987**). ''A hierarchical neural-network model for control and learning of voluntary movement,'' Biol. Cybern. **57**, 169–185.

Keele, S. W. (**1968**). ''Movement control in skilled motor performance,'' Psychol. Bull. **70**, 387–403.

Kirchner, J. A., and Wyke, B. D. (**1965**). ''Articular reflex mechanisms in the larynx,'' Ann. Otol. Rhinol. Laryngol. **74**, 748–768.

Klatt, D. (**1973**). ''Discrimination of fundamental frequency contours in synthetic speech: Implications for models of pitch perception,'' J. Acoust. Soc. Am. **53**, 8–16.

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

Larson, C. R. (**1988**). ''Brain mechanisms involved in the control of vocalization,'' J. Voice **2**, 301–311.

Larson, C. R., Burnett, T. A., Kiran, S., and Hain, T. C. (**2000**). ''Effects of pitch-shift velocity on voice $F0$ responses,'' J. Acoust. Soc. Am. **107**, 559–564.

McFarland, D. H., and Baum, S. R. (**1995**). ''Incomplete compensation to articulatory perturbation,'' J. Acoust. Soc. Am. **97**, 1865–1873.

Miall, R. C., and Wolpert, D. M. (**1996**). ''Forward models for physiological motor control,'' Neural Networks **9**, 1265–1279.

Milenkovic, P., and Read, C. (**1995**). *CSpeechSP Quick Reference Manual*, University of Wisconsin.

Milner, A. D., and Goodale, M. A. (**1995**). *The Visual Brain in Action* (Oxford University Press, New York).

Perkell, J., 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**, 2961–2979.

Perkell, J., Matthies, M., Lane, H., Guenther, F., Wilhelms-Tricarico, R., Wozniak, J., and Guoid, P. (**1997**). ''Speech motor control: Acoustic goals, saturation effects, auditory feedback and internal models,'' Speech Commun. **22**, 227–250.

Sapir, S., McClean, M. D., and Larson, C. R. (**1983**). ''Human laryngeal responses to auditory stimulation,'' J. Acoust. Soc. Am. **73**, 315–321.

Shadmehr, R., and Mussa-Ivaldi, F. A. (**1994**). ''Adaptive representation of dynamics during learning of a motor task,'' J. Neurosci. **14**, 3208–3224.

Titze, I. R. (**1994**). *Principles of Voice Production* (Prentice-Hall, Englewood Cliffs, NJ).

von Holst, E., and Mittelstaedt, H. (**1950**). *The principle of reafference: Interactions between the central nervous system and the peripheral organs*, edited by P. C. Dodwell (Appleton-Century-Crofts, New York, 1971).

Welch, R. B. (**1986**). ''Adaptation of space perception,'' in *Handbook of Perception and Human Performance*, edited by K. R. Boff, L. Kaufman, and J. P. Thomas (Wiley, New York).

Zemlin, W. R. (**1981**). *Speech and Hearing Science: Anatomy and Physiology*, 2nd ed. (Prentice-Hall, Englewood Cliffs, NJ).