Behavioral/Systems/Cognitive

Simultaneous Acquisition of Multiple Auditory–Motor Transformations in Speech

Amélie Rochet-Capellan[1] **and David J. Ostry**[1,2]

[1]Department of Psychology, McGill University, Montreal, Quebec, Canada H3A 1B1, and [2]Haskins Laboratories, New Haven, Connecticut 06511

The brain easily generates the movement that is needed in a given situation. Yet surprisingly, the results of experimental studies suggest that it is difficult to acquire more than one skill at a time. To do so, it has generally been necessary to link the required movement to arbitrary cues. In the present study, we show that speech motor learning provides an informative model for the acquisition of multiple sensorimotor skills. During training, subjects were required to repeat aloud individual words in random order while auditory feedback was altered in real-time in different ways for the different words. We found that subjects can quite readily and simultaneously modify their speech movements to correct for these different auditory transformations. This multiple learning occurs effortlessly without explicit cues and without any apparent awareness of the perturbation. The ability to simultaneously learn several different auditory–motor transformations is consistent with the idea that, in speech motor learning, the brain acquires instance-specific memories. The results support the hypothesis that speech motor learning is fundamentally local.

Introduction

Does sensorimotor learning result in representations for motor control that generalize broadly to new targets and situations, or does it result in local skills, specific to the target and the situation of the learning? Patterns of generalization of motor learning in both speech production and limb movement suggest that learning is largely local. In contrast, the difficulty in work on human arm movement to learn multiple sensorimotor adaptations at a given place and time suggests that motor learning is not easily fractionated. But does the difficulty observed in arm movement reveal a fundamental limit on learning or is it linked to the use of arm movement as a model system for these tests? In the present paper, we show that simultaneous sensorimotor acquisitions are readily observable in speech. Speech data are thus consistent with the idea that sensorimotor learning is local.

The rather specific nature of motor learning is indicated by the observation that learning only generalizes to a limited extent for movements not experienced during training. In arm movements, adaptation to force fields or visuomotor perturbations transfers poorly to movements in directions other than those in the training direction (Gandolfo et al., 1996; Ghahramani et al., 1996; Krakauer et al., 2000; Witney and Wolpert, 2003; Mattar and Ostry, 2007). There is likewise lim-

ited transfer to untrained movement amplitudes (Mattar and Ostry, 2010) and little transfer of force-field adaptation between limbs (Malfait and Ostry, 2004). Similarly, speech adaptation to mechanical perturbations of the jaw trajectory transfers poorly to the subsequent production of nonspeech movements (Tremblay et al., 2003) and to speech material not used in training (Tremblay et al., 2008).

This pattern of largely local effects can be contrasted with the actual learning process where, at least in the case of arm movement, there is weaker evidence for specificity. In studies of arm movement, subjects have difficulty learning more than one sensorimotor transformation at a time (Gupta and Ashe, 2007). To do so, it has been necessary to associate the perturbations with different limbs (Bock et al., 2005; Galea and Miall, 2006) or different locations in space (Woolley et al., 2007), to introduce a delay between training and testing phases (Brashers-Krug et al., 1996), or to provide subjects with arbitrary visual cues and/or explicit contextual information (Wada et al., 2003; Osu et al., 2004; Imamizu et al., 2007).

Similar difficulties in learning multiple sensorimotor transformations in parallel may not occur in other model systems. In speech learning, in particular, sounds and associated movements are acquired in a manner that is specific to the word in which the sound is embedded (Tremblay et al., 2008). Accordingly, we have developed a multiple-adaptation paradigm for speech motor learning to assess the extent to which subjects can simultaneously adapt to several different auditory–motor transformations. We created an experimental model of the speech learning process in which subjects receive auditory feedback that is altered in opposing ways for different vowels in different words (experiments 1 and 2) and to the same vowel in different words (experiments 3 and 4). We show that subjects readily and spontaneously corrected these op-

posing perturbations. The results are consistent with the idea of instance-specific learning of speech movements.

Materials and Methods

Subjects and task. We tested sixty-five native speakers of English that had no reported impairment of hearing or speech. There were 12 females and seven males in experiments 1 (mean age, 21.3 ± 2.0 years) and 2 (mean age, 21.7 ± 2.7), nine females and five males in experiment 3 (mean age, 20.6 ± 1.4 years), and 11 females and two males in experiment 4 (mean age, 20.1 ± 2.1 years). All participants signed consent forms approved by McGill University Institutional Review Board.

The task was to read words aloud that were displayed one at a time on a computer monitor. Subjects were told that they would have to speak into a microphone and that they would hear their own voice mixed with noise through earphones. They were also asked to speak quietly to avoid receiving auditory feedback other than through the headphones. Subjects were instructed to use a normal duration for all words and to stay ~15 cm from the microphone. The experimenter did not indicate that the auditory feedback would be altered.

Following the experiment, subjects were asked whether they had noticed anything special about the auditory feedback. None reported that the sound of the words had changed or that one word sounded more like another. Most remarks about auditory feedback were about the volume. Subjects did not have the feeling that they had changed their speech.

Test words and auditory perturbation. Vowel sounds are characterized by vocal tract resonances known as formants that differ in frequency for different vowels. The first two formants (F1 and F2) contain most of the acoustical energy and are most important in distinguishing between vowels. By changing formant frequencies in the acoustical signal, it is possible to make one vowel more or less acoustically similar to another vowel (Houde and Jordan, 1998; Purcell and Munhall, 2006). This is why we used vowel sounds rather than consonants in these experiments. We chose the vowel /ɛ/ for these studies, because this sound can readily be transformed so as to sound more like /æ/ in "had" (by increasing the F1 frequency) or more like /I/ in "hid" (by decreasing the F1 frequency) (Fig. 1A).

In experiments 1 and 2, we evaluated subjects' ability to compensate for opposing F1 shifts to the vowels /æ/ in "had" and /ɛ/ in "head." The applied perturbations either decreased the F1 distance between the two words, so that they sounded more similar (experiment 1) or increased the F1 distance, which drew the two words apart in terms of auditory feedback (experiment 2) (Fig. 1B1,B2). In experiment 3, we evaluated subjects' ability to compensate for opposing perturbations applied to the vowel /ɛ/ in "bed," "head," and "ted." F1 was increased in "bed," decreased in "head," and not altered in "ted" (Fig. 1B3). Experiment 4 served as a control for experiment 3; F1 was increased in both "bed" and "head" and was not altered in "ted" (Fig. 1B4). This let us test whether the ability to maintain F1 in "ted" in experiment 3 resulted from an averaging of the opposing adaptations for "head" and "ted" or from the ability to maintain each target separately.

Experimental procedures. Subjects were seated at a table in a soundproof room in front of a computer monitor (Fig. 2). They wore earphones (SR001-MK2 electrostatic; Stax) throughout the entire experiment. The test words were displayed in random order, one word at a time for 1.2 s; successive words were presented 1.2 s apart. Before starting the experiment, subjects were familiarized with the basic procedure, which involved speaking and hearing their amplified but unshifted voice mixed with the noise through the earphones.
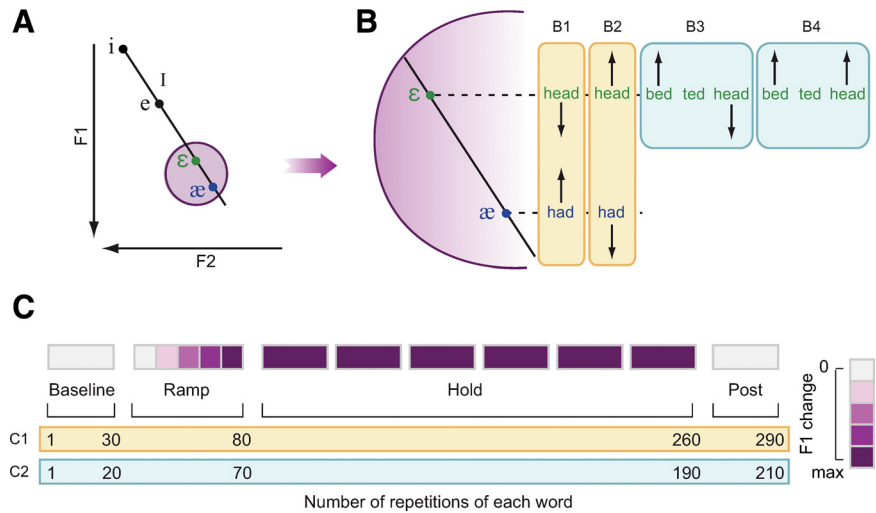

Figure 1. Experimental procedure. *A*, Front vowels used in the study (simplified representation) (Ladefoged, 2001). *B*, F1 shift for different experimental groups. *B1*, Experiment 1. F1 in "head" was increased toward "had" and F1 in "had" was decreased toward "head." *B2*, Experiment 2. F1 in "head" was decreased and F1 in "had" was increased. *B3*, Experiment 3. F1 was increased for "bed," decreased for "head," and unaltered for "ted." *B4*, Experiment 4. F1 was increased in both "head" and "bed" and unaltered for "ted." Arrows on the F1 and F2 axes show the direction of F1/F2 increase. *C*, Experimental sequence. Each session included the same number of productions of each word in random order. After an initial block with unaltered feedback, the perturbation of the auditory feedback was introduced in five steps, maintained for six blocks, and returned to normal in the final block. Blanks between blocks indicate short breaks. The total number of repetitions of each word was 290 in experiments 1 and 2 (*C1*) and 210 in experiments 3 and 4 (*C2*).

The experiment consisted of nine short blocks of word repetition separated by 1–2 min pauses (Fig. 1C). Each block contained an equal number of repetitions of each word, mixed in random order. After an initial block with normal feedback [30 trials per word in experiments 1 and 2 and 20 trials per word in experiments 3 and 4 (Fig. 1C1,C2)], the shift was introduced gradually in five discrete steps. Each step included 10 randomized repetitions of each word. The shift was then held at a maximum value for six blocks of 30 trials per word in experiments 1 and 2 and 20 trials per word in experiments 3 and 4. Feedback was returned to normal (unshifted) in a final block of 30 repetitions of each word in experiment 1 and 2 and 20 repetitions in experiments 3 and 4.

The subjects' voices were recorded using a unidirectional microphone (Sennheiser) and digitally sampled at 44,100 Hz. In parallel to the recording, the acoustical signal was processed in real time and played back to participants (Fig. 2). After preamplification (with a level that was set separately for each subject), the signal from the microphone was split into two paths. In the low-frequency path, an electronic speech processor (VoiceOne; TC Helicon) shifted all frequencies of the original signal except the pitch. The mean downward shift in F1 was 85% ± 0.5% (mean ± SE) of the initial value for both the vowels /ɛ/ and /æ/. The mean upward shift was 125% ± 0.8% of the initial F1 value for the vowel /ɛ/ and 120% ± 0.9% for the vowel /æ/. After being shifted, the signal was then analog low-pass filtered with a frequency that was set on a per subject basis to preserve the pitch and the shifted first-formant frequency while discarding all higher frequencies. In the high-frequency path, the signal was electronically delayed by 11 ms. This delay compensated for an equal delay introduced by the VoiceOne and the low-pass filter. The signal was then analog high-pass filtered at the same frequency as that used in the low-frequency path. This process excluded the fundamental frequency and the first formant but preserved the unshifted higher formant frequencies. The signals from the two paths were then mixed together and the resulting signal was amplified and played back to participants.

Two methods were used to limit the reception of the unshifted signal that might be present because of bone conduction or the airborne signal outside the earphones. First, speech-shaped masking noise was mixed with the processed signal presented to subjects. The volume of the noise was the same for all the subjects (70 dB through the headphones, as measured by a sound-level meter). Second, the volume of the speech
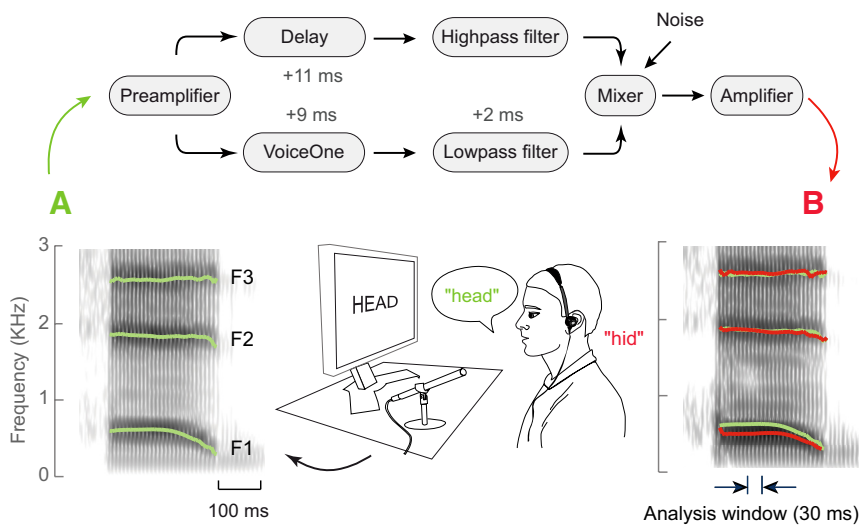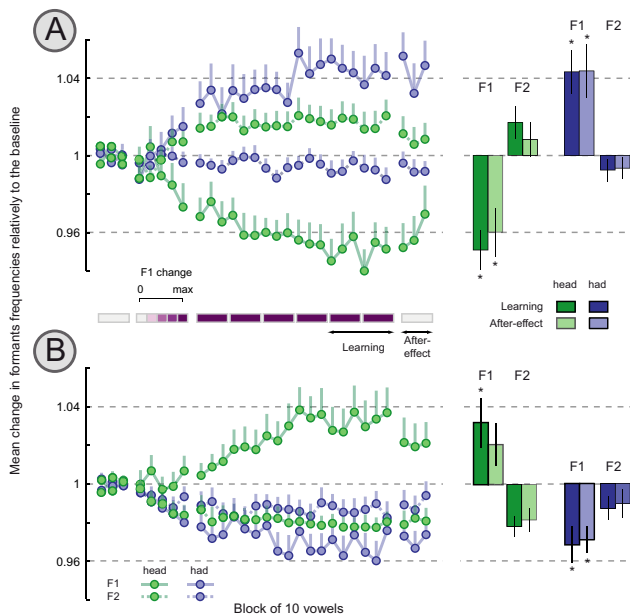
Figure 2. Experimental set-up. The subject read words aloud and heard his altered voice through open earphones. The sound pattern of "head" (spectrogram shown in **A**, formant tracking in green) was altered to sound like "hid" (F1 shifted downward, spectrogram shown in **B**, formant tracking in red). The selective shift of F1 was accomplished by using an acoustical effects processor in combination with filters and mixing devices.

Figure 3. **A**, **B**, Average values of F1 and F2 relative to frequencies in prelearning trials for experiments 1 (**A**, first row) and 2 (**B**, second row). Left, Curves give the progression of frequency change over the course of learning. Right, Bar graphs show means and the SEs for learning and aftereffects. Asterisks indicate reliable differences ($p < 0.05$). **A**, In experiment 1, F1 auditory feedback was shifted down for "had" and up for "head." In response to the perturbation, subjects increased F1 values for "had" and decreased F1 for "head." **B**, In experiment 2, a change in F1 opposite to that in experiment 1 resulted in opposing compensations. In both experiments, F2 tended to change in a direction opposite to F1 for "head" whereas no clear change of F2 was observed for "had."

signal in the earphones was set to be louder than normal, but without creating discomfort, masking the unshifted signal and leading the subject to speak more softly.

Data analysis. Following the experiment, the acoustic signals were resampled at 22,050 Hz in experiments 1 and 2 and 10,000 Hz in experiments 3 and 4. The boundaries of vowels were determined automatically using the Praat software environment (freeware provided by Paul Boersma and David Weenink, Phonetic Sciences, University of Amsterdam, Amsterdam, The Netherlands), then manually

checked and, if necessary, corrected. The speech signals were verified by listening; trials with errors of production or noise were discarded from the analyses.

We analyzed both F1 and F2 values, as previous studies have shown that adaptation to F1 shifts may also result in small changes to F2 (Villacorta et al., 2007). Praat was used to compute the two first formants for each utterance, based on a window of 30 ms in the center of the vowel (Fig. 2). The range of frequencies for the detection of the formants was adapted to the subject and chosen to reduce the variability of F1 across the recorded productions. Automatic detection of F1 was successful for all subjects. Detection of F2 frequencies failed for three subjects in experiment 1 and five subjects in experiment 2. For these eight subjects, F2 was corrected manually based on the spectrogram. We discarded trials for which F1 or F2 were beyond ±2 SDs from the mean of each phase with the same perturbation value (each step of the ramp phase).

Motor learning was evaluated separately for F1 and F2 by comparing the mean of frequency values during the baseline phase (first block of repetitions) (Fig. 1C) with the mean of frequencies during the two last blocks of the hold phase (blocks 7 and 8). Aftereffects were evaluated in the same way by comparing the baseline phase with the final block of the experiment (block 9). Statistical reliability of adaptation and aftereffects were evaluated for each experiment and each formant separately using within-subjects ANOVA with two factors: the experimental utterance ("head" vs "had" in experiment 1 and 2; and "head" vs "ted" vs "bed" in experiments 3 and 4) and the phase of the experimental manipulation. Bonferroni-corrected comparisons were conducted to test individual contrasts.

Results

In experiment 1, the auditory feedback of F1 was increased for "head" and decreased for "had." As shown in Figure 3A, subjects compensated for these perturbations by modifying F1 in their productions downward in "head" and upward in "had." The perturbation of F1 also induced a change of F2 in a direction opposite to F1 for "head" utterances; there was no clear change in F2 in "had" utterances. When normal feedback was restored at the end of the experiment, we observed aftereffects that were in the same direction as the compensation.

ANOVA showed that changes to F1 in different phases of the experiment depended on the experimental utterance ($F_{(2,36)} = 32.8$, $p < 0.0001$). F1 decreased over the course of training by 32.4 Hz for "head" ($p < 0.001$) and increased by 36.3 Hz for "had" ($p < 0.01$). This corresponds to a change of -5% and $+4.3\%$, respectively. When normal feedback was restored at the end of the experiment, F1 remained 27.7 Hz (4.0%) lower and 36.3 Hz (4.4%) higher than during the baseline trials for "head" ($p < 0.05$) and "had" ($p < 0.01$), respectively. Changes to F1 were accompanied by small changes in F2 in the opposite direction, which had different patterns over the course of the experiment for different test words ($F_{(2,36)} = 7.7$, $p < 0.01$). At the end of training, F2 decreased for "head" and increased for "had," but none of the *post hoc* comparisons for F2 was statistically reliable.

In a second experiment, we assessed whether these multiple adaptations were also present in response to opposing perturbations that moved the sounds of "head" and "had" apart in terms of auditory feedback. As shown in the Figure 3B, in response to these perturbations, subjects progressively increased F1 in their

"head" utterances and decreased F1 in their "had" utterances. As in experiment 1, ANOVA showed that the compensatory frequency shifts for the two words depended on the experimental phase ($F_{(2,36)}$ = 19.0, $p < 0.0001$). At the end of the training phase, F1 increased by 23.1 Hz for "head" ($p < 0.05$) and decreased by 25.7 Hz for "had" ($p < 0.01$), which corresponds to a relative change of $+3.2\%$ and -3.2%, respectively. When normal feedback was restored, F1 remained, on average, 24.0 Hz (3%) smaller than the baseline mean for "had" ($p < 0.01$). A $+14.2$ Hz (2%) aftereffect for "head" was not reliable ($p > 0.15$).

As in the first experiment, changes in F1 were accompanied by small changes in F2. Unlike experiment 1, the F2 changes at the end of the training were in the same direction for both "head" (-40.6 Hz) and "had" (-21.3 Hz). There were also F2 aftereffects (-34.2 Hz for "head" and -17.7 Hz for "had"). ANOVA indicated reliable differences in F2 over the course of the experiment ($F_{(2,36)}$ = 8.7, $p < 0.01$). These changes were similar for the two words. Both words together had F2 values that significantly differed between the baseline and the end of the training ($p < 0.001$) and between the baseline and the post-phase ($p < 0.05$). The statistical interaction between test words and experimental phase was not reliable ($F_{(2,36)}$ = 1.7, $p > 0.2$).

In experiment 3, we tested whether subjects could learn to produce the same vowel differently depending on the word in which it was embedded. As shown in Figure 4A, at the end of the training, subjects modified F1 in their "head" and "bed" productions in directions opposite to the imposed perturbation. Aftereffects were observed when the frequency shift was removed. There were no systematic changes for "ted", which was produced with unshifted auditory feedback.

ANOVA indicated that the pattern of F1 differed for the three words in a manner that depended on the phase of the experiment ($F_{(4,52)}$ = 36.5 $p < 0.0001$). For "bed" utterances, which were produced with auditory feedback shifted upward, F1 was 43.3 Hz (6.7%) lower than baseline in the two final training blocks ($p < 0.001$) and 27.8 Hz (4.3%) lower in the aftereffect block ($p < 0.01$). For "head", where auditory feedback was shifted downward, F1 increased by 42.4 Hz (5.8%) at the end of training ($p < 0.05$) and was still 34.0 Hz (4.8%) greater than baseline during aftereffect trials ($p < 0.05$). For "ted", which was produced with unshifted auditory feedback, F1 did not differ significantly from baseline values either at the end of training or during the aftereffect phase ($+1.7\%$ and $+1.9\%$, $p > 0.4$ in both cases).

As observed for "head" utterances in experiments 1 and 2, F2 changed in a direction opposite to F1 for both "head" and "bed." As before, ANOVA indicated that the F2 pattern for different words reliably varied according to the phase of the experiment ($F_{(4,52)}$ = 20.3, $p < 0.0001$). The measured change from baseline to end of training was $+47.9$ Hz (2.5%) for "bed" ($p < 0.05$) and -19.7 Hz (1.2%) for "head" ($p > 0.3$). In contrast, F2 changed in the same direction as F1 for "ted" ($+15.3$ Hz, 0.1%) but this
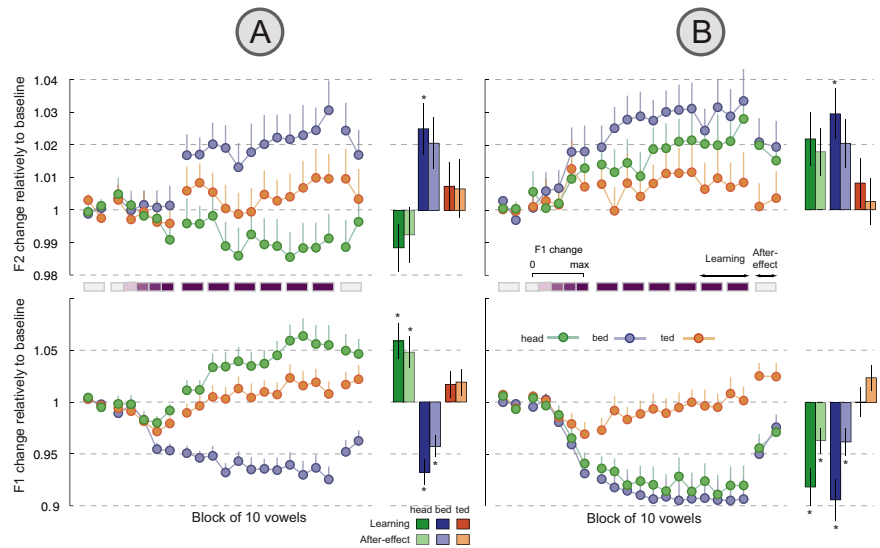
Figure 4. *A*, *B*, Average values of F1 and F2 relative to prelearning frequencies in experiments 3 (*A*, left) and 4 (*B*, right). For each experiment, the curves on the left represent the progression of frequency changes over the courses of learning for F1 (bottom) and F2 (top). Right, Bar graphs show the means and the SEs for learning and aftereffects. Asterisks indicate reliable effects ($p < 0.05$). *A*, In experiment 3, F1 auditory feedback was increased for "bed," decreased for "head," and unaltered for "ted." In response to these perturbations, the subjects progressively decreased F1 in "bed" and increased F1 in "head." *B*, In experiment 4, F1 was increased in both "head" and "bed" and unaltered for "ted." Subjects compensated for the perturbation by progressively decreasing F1 in both "bed" and "head." In both experiments, F2 values tended to change in a direction opposite to F1 for the shifted words. No systematic modification of F1 or F2 appeared for the unshifted frequencies in "ted."

change was not reliable ($p > 0.9$). Aftereffects in F2 appeared to be present following the removal of altered auditory feedback but none of the *post hoc* comparisons for F2 aftereffects was statistically reliable ($p > 0.06$ in all cases).

In experiment 4, we evaluated whether unchanged productions of "ted" reflected an averaging of opposing adaptations for "head" and "bed" or the ability to independently maintain F1 frequency in "ted." F1 frequency was shifted upward for both "head" and "bed" and unshifted for "ted." As shown in Figure 4B, with training, subjects lowered F1 and raised F2 in both "head" and "bed" productions in a similar fashion. F1 and F2 frequencies both differed significantly over the course of the experiment in a manner that depended on the training utterance ($F_{(4,48)}$ = 36.5, $p < 0.0001$ for F1; $F_{(4,48)}$ = 8.9, $p < 0.0001$ for F2). At the end of the training phase, F1 decreased relative to baseline values by 56.6 Hz for "head" (-8%, $p < 0.01$) and by 67.2 Hz for "bed" (-9%, $p < 0.01$). F1 changes were not observed for "ted" ($+0.5$ Hz, $p > 0.9$). In parallel with the F1 decrease, F2 increased by 43.6 Hz for "head" ($+2.2\%$) and by 58.2 Hz for "bed" ($+2.9\%$). Only the change in bed was reliable ($p < 0.01$). As in experiment 3, F2 changed in the same direction as F1 for the unshifted word "ted" ($+16.2$ Hz, $+0.5\%$), but the change was not reliable ($p > 0.9$). Both F1 and F2 had aftereffects in the same direction as the adaptation, with F1 values lower than those under baseline conditions and F2 values higher. During the aftereffect phase, F1 was significantly different from baseline for both "head" and "bed" ($p < 0.05$). The increase in F1 observed in "ted" during the post-phase was not significant ($p > 0.1$). Aftereffects were not reliable for F2 ($p > 0.05$ in all cases).

Finally, we assessed whether the magnitude of adaptation for an individual utterance was affected by co-occurring adaptation to other words. We focused on "bed," which was shifted in the same direction in experiments 3 and 4, and

"head," which was shifted in the same direction in experiments 2 and 3. For "bed," we found no significant differences in the magnitude of the F1 ($F_{(1,25)} = 1.3$, $p > 0.2$) or F2 ($F_{(1,25)} = 0.2$, $p > 0.6$) change in the two experiments. Similarly, for "head", there were no significant differences in F1 and F2 changes caused by learning ($F_{(1,31)} = 1.56$, $p > 0.2$, $F_{(1,31)} = 1.35$, $p > 0.25$, respectively). Thus, adaptation magnitude for individual items does not appear to be affected by concurrent adaptation to other words.

In summary, in experiments 1 and 2, subjects learned to compensate for opposing changes to F1 in "head" and "had" by modifying F1 in their productions in opposing directions for the two words. Experiments 3 and 4 demonstrated that multiple adaptations are also observed when different perturbations are applied to the same vowel in different words.

Discussion

The present studies show that subjects can simultaneously modify speech movements in different ways for different vowels (experiment 1 and 2) and when the same vowel is embedded in different words (experiments 3 and 4) . The ease with which the present adaptations are achieved in speech can be contrasted with the difficulty observed in achieving multiple simultaneous adaptations in arm movement studies. The multiple adaptations observed here suggest that speakers adapt their motor commands specifically to the target utterance, which points to the specificity of speech motor learning.

There are several pieces of evidence that show specificity in speech motor learning in the present study. First, when speakers encounter opposing auditory transformations, they are able to compensate in parallel for each transformation. This means that they separately change their motor commands to correct for the different experimentally imposed perturbations. We see this ability in experiments 1 and 2 when different vowels are transformed in different directions and in experiment 3 when the same vowel embedded in different words is transformed in different directions. A second piece of evidence for specificity in learning is the fact that subjects are able to change their vocal output for some words and not for others. We see this ability in experiments 3 and 4, where two words are shifted and one is maintained. A final piece of evidence for specificity is the magnitude of the compensation. In experiments 2 and 3 (for "head") and experiments 3 and 4 (for "bed"), we find that, for a given utterance and direction of formant shift, the amount of compensation does not depend on transformations applied to other utterances in the experiment. Thus, learning for one utterance appears to interfere little with concurrent learning for another utterance.

The properties of the multiple adaptations seen here are consistent with the characteristics of auditory–motor adaptations observed previously in the context of individual auditory perturbations (Houde and Jordan, 1998; Purcell and Munhall, 2006; Villacorta et al., 2007). The compensation occurred progressively over the course of the training, reached an asymptotic level, and persisted in the form of aftereffects when the auditory feedback was returned to normal following training. As in the case of single adaptations, the compensation observed here primarily involved changes to F1, but there were also opposing changes to F2, especially for the vowel /ε/. The changes observed in F2 may be due to the fact that front vowels depend both on the F1 and F2 values (Ladefoged, 2001) (Fig. 1A).

The results of the present studies show that subjects can acquire different sensorimotor transformations at the same place and time. The results are inconsistent with the hypothesis that sensorimotor adaptation in speech occurs by altering a global predictive model that is shared by the production control machinery for different movements (Houde and Jordan, 2002; Villacorta et al., 2007; Bohland et al., 2010). Indeed, if motor control in speech involved a general predictive representation, subjects should not have been able to compensate in different ways for different transformations. Instead, the ability to produce different patterns of adaptation for different utterances suggests that the learning observed here is local or instance-specific. This result is consistent with previous work that shows little transfer of sensorimotor learning either in speech (Tremblay et al., 2008) or in human arm movement (Gandolfo et al., 1996; Ghahramani et al., 1996; Krakauer et al., 2000; Witney and Wolpert, 2003; Malfait and Ostry, 2004; Mattar and Ostry, 2007, 2010).

The present studies, in which subjects simultaneously adapt to different auditory transformations to different words should be contrasted with previous studies of adaptation to altered F1 feedback that have shown trial-to-trial transfer over the course of learning (Houde and Jordan, 2002; Villacorta et al., 2007). The latter demonstrations, conducted using masked auditory feedback, have been interpreted as evidence for a global mechanism dedicated to vowel production (Houde and Jordan, 2002; Villacorta et al., 2007). If a global mechanism such as this were to exist, then subjects should not have been able to adapt their movements in different ways for the production of vowels as we observed here. Subjects' ability to modify motor commands for the production of the same vowel in different words suggests that vowel production is tied specifically to the words in which they are embedded. In this sense, the present experiments have bearing on the question of the nature of the representations and processes that underlie speech learning and production.

References

Bock O, Worringham C, Thomas M (2005) Concurrent adaptations of left and right arms to opposite visual distortions. Exp Brain Res 162:513–519.

Bohland JW, Bullock D, Guenther FH (2010) Neural representations and mechanisms for the performance of simple speech sequences. J Cogn Neurosci 22:1504–1529.

Brashers-Krug T, Shadmehr R, Bizzi E (1996) Consolidation in human motor memory. Nature 382:252–255.

Galea JM, Miall RC (2006) Concurrent adaptation to opposing visual displacements during an alternating movement. Exp Brain Res 175:676–688.

Gandolfo F, Mussa-Ivaldi FA, Bizzi E (1996) Motor learning by field approximation. Proc Natl Acad Sci U S A 93:3843–3846.

Ghahramani Z, Wolpert DM, Jordan MI (1996) Generalization to local remappings of the visuomotor coordinate transformation. J Neurosci 16:7085–7096.

Gupta R, Ashe J (2007) Lack of adaptation to random conflicting force fields of variable magnitude. J Neurophysiol 97:738–745.

Houde JF, Jordan MI (1998) Sensorimotor adaptation in speech production. Science 279:1213–1216.

Houde JF, Jordan MI (2002) Sensorimotor adaptation of speech. I. Compensation and adaptation. J Speech Lang Hear Res 45:295–310.

Imamizu H, Sugimoto N, Osu R, Tsutsui K, Sugiyama K, Wada Y, Kawato M (2007) Explicit contextual information selectively contributes to predictive switching of internal models. Exp Brain Res 181:395–408.

Krakauer JW, Pine ZM, Ghilardi MF, Ghez C (2000) Learning of visuomotor transformations for vectorial planning of reaching trajectories. J Neurosci 20:8916–8924.

Ladefoged P (2001) Vowels and consonants: an introduction to the sounds of languages. Oxford: Blackwell Publishers.

Malfait N, Ostry DJ (2004) Is interlimb transfer of force-field adaptation a "cognitive" response to the sudden introduction of load? J Neurosci 24:8084–8089.

Mattar AA, Ostry DJ (2007) Neural averaging in motor learning. J Neurophysiol 97:220–228.

Mattar AA, Ostry DJ (2010) Generalization of dynamics learning across changes in movement amplitude. J Neurophysiol 104:426–438.

Osu R, Hirai S, Yoshioka T, Kawato M (2004) Random presentation enables subjects to adapt to two opposing forces on the hand. Nat Neurosci 7:111–112.

Purcell DW, Munhall KG (2006) Adaptive control of vowel formant frequency: Evidence from real-time formant manipulation. J Acoust Soc Am 120:966–977.

Tremblay S, Shiller DM, Ostry DJ (2003) Somatosensory basis of speech production. Nature 423:866–869.

Tremblay S, Houle G, Ostry DJ (2008) Specificity of speech motor learning. J Neurosci 28:2426–2434.

Villacorta VM, Perkell JS, Guenther FH (2007) Sensorimotor adaptation to feedback perturbations of vowel acoustics and its relation to perception. J Acoust Soc Am 122:2306–2319.

Wada Y, Kawabata Y, Kotosaka S, Yamamoto K, Kitazawa S, Kawato M (2003) Acquisition and contextual switching of multiple internal models for different viscous force fields. Neurosci Res 46:319–331.

Witney AG, Wolpert DM (2003) Spatial representation of predictive motor learning. J Neurophysiol 89:1837–1843.

Woolley DG, Tresilian JR, Carson RG, Riek S (2007) Dual adaptation to two opposing visuomotor rotations when each is associated with different regions of workspace. Exp Brain Res 179:155–165.