J Neurophysiol 107: 1711–1717, 2012.
First published December 21, 2011; doi:10.1152/jn.00773.2011.

Nonhomogeneous transfer reveals specificity in speech motor learning

Amélie Rochet-Capellan,[1] Lara Richer,[1] and David J. Ostry[1,2]
[1]*McGill University, Montreal, Canada; and* [2]*Haskins Laboratories, New Haven, Connecticut*

Submitted 23 August 2011; accepted in final form 14 December 2011

Rochet-Capellan A, Richer L, Ostry DJ. Nonhomogeneous transfer reveals specificity in speech motor learning. *J Neurophysiol* 107: 1711–1717, 2012. First published December 21, 2011; doi:10.1152/jn.00773.2011.—Does motor learning generalize to new situations that are not experienced during training, or is motor learning essentially specific to the training situation? In the present experiments, we use speech production as a model to investigate generalization in motor learning. We tested for generalization from training to transfer utterances by varying the acoustical similarity between these two sets of utterances. During the training phase of the experiment, subjects received auditory feedback that was altered in real time as they repeated a single consonant-vowel-consonant utterance. Different groups of subjects were trained with different consonant-vowel-consonant utterances, which differed from a subsequent transfer utterance in terms of the initial consonant or vowel. During the adaptation phase of the experiment, we observed that subjects in all groups progressively changed their speech output to compensate for the perturbation (altered auditory feedback). After learning, we tested for generalization by having all subjects produce the same single transfer utterance while receiving unaltered auditory feedback. We observed limited transfer of learning, which depended on the acoustical similarity between the training and the transfer utterances. The gradients of generalization observed here are comparable to those observed in limb movement. The present findings are consistent with the conclusion that speech learning remains specific to individual instances of learning.

sensorimotor learning; speech production; generalization; specificity

THE ABILITY TO APPLY MOTOR skills to novel situations is an indicator of generalization or transfer of learning. The analysis of generalization in motor learning is a behavioral window into the way that the brain uses past experiences to act in new situations. Generalization observed under controlled laboratory situations also serves as a model for rehabilitation protocols (Bastian 2010; Maas et al. 2008) and provides insights into the processes that subtend the control of movements (Gandolfo et al. 1996; Houde and Jordan 1998; Shadmehr 2004; Wolpert et al. 1995). Generalization has been extensively investigated in the arm movement literature. The findings show that generalization depends on the amount of overlap between the movements experienced in the course of training and those involved in the assessment of transfer. In contrast, few empirical studies have addressed generalization in speech motor learning. In the present paper, we examine transfer of speech motor learning in adults following adaptation to auditory feedback that is altered in real time. We find that, despite substantial adaptation to individual training utterances, speech learning transfers poorly from one sound to another. The magnitude of transfer is seen to vary in a systematic fashion with the distance in sound space between the training and transfer utterances. These results suggest that speech motor learning is local, or specific, to the training material.

Local learning as it relates to motor function is the idea that a unique configuration of motor commands is acquired and maintained to produce individual movements (Atkeson 1989). Local learning is indicated by the presence of tuning curves that show a progressive reduction in transfer of learning as the difference between training and transfer conditions grows. Learning in the context of arm movement is characterized by such a graded pattern of generalization that depends on the overlap between the properties of the training experience and those of the transfer task. Hence, motor learning in one direction transfers to movements in other directions as a function of the angular distance between the two directions (Donchin et al. 2003; Gandolfo et al. 1996; Ghahramani and Wolpert 1997; Krakauer et al. 2000; Mattar and Ostry 2007b; Thoroughman and Shadmehr 2000; Thoroughman and Taylor 2005). Similarly, generalization of force-field learning to movements of different amplitudes occurs only when the amplitude of the training movement includes the amplitude of the transfer movement (Mattar and Ostry 2010). Learning is also seen to be linked with the presence of implicit or explicit contexts or cues (Cothros et al. 2009; Imamizu et al. 2007; Krakauer et al. 2006; Osu et al. 2004; Wada et al. 2003). When generalization of learning is observed, it is typically dependent on an interpolation of past local learning experiences (Gandolfo et al. 1996; Ghahramani and Wolpert 1997; Malfait et al. 2005; Mattar and Ostry 2007a).

Generalization of speech motor learning is fundamental, both for the understanding of the relationships between motor and linguistics aspects of language production (Houde and Jordan 1998), and likewise for the development of rehabilitation protocols. Yet few studies have addressed the properties of generalization in speech learning. When training and transfer movements are segregated, adaptation to mechanical loads does not transfer to untrained utterances, nor is there transfer to nonspeech orofacial movements (Tremblay et al. 2003; Tremblay et al. 2008). Speakers are also able to learn several auditory-motor transformations in parallel, which supports the notion that speech learning is local (Rochet-Capellan and Ostry 2011). Generalization from one sound to another has been observed when transfer is tested over the course of learning and when several training utterances are mixed with several transfer utterances (Cai et al. 2010; Houde and Jordan 1998; Villacorta et al. 2007). However, under these conditions, the patterns of generalization observed are difficult to interpret, as transfer could reflect an averaging that takes places when subjects experience several training conditions simultaneously (Mattar and Ostry 2010; Takahashi et al. 2001).

The present paper addresses generalization in speech motor learning using a procedure in which training and test trials are

Address for reprint requests and other correspondence: D. J. Ostry, Dept. of Psychology, McGill Univ., 1205 Dr. Penfield Ave. Stewart Biology Bldg, Montreal, Quebec, Canada H3A 1B1 (e-mail: david.ostry@mcgill.ca).

presented in separated blocks, with different groups of subjects each tested with a single training utterance and a single transfer utterance. We studied motor learning using an auditory-motor transformation. Subjects were required to repeat aloud individual utterances, while the frequency composition of their auditory feedback was altered and played back to them in real-time through headphones (Houde and Jordan 1998; Purcell and Munhall 2006; Rochet-Capellan and Ostry 2011; Villacorta et al. 2007). We observed that subjects progressively learned to compensate for the auditory perturbation. Following learning, we tested for transfer by having subjects repeat a transfer utterance that differed from the training utterance in terms of its distance in sound space. We found that speech motor learning generalized poorly and that the amount of generalization varied with the distance in sound space between training and transfer utterances. Our results are consistent with previous work on human limb movement and speech and suggest that motor learning is fundamentally local.

METHODS

Subjects. Subjects were native speakers of English between the ages of 18 and 30 yr and had no reported impairment of hearing or speech. All participants signed consent forms approved by the McGill University Institutional Review Board.

Word repetition with altered auditory feedback. The subjects' task was to read aloud words that were displayed one at a time on a computer monitor. Auditory feedback was provided through headphones. Subjects were informed they would hear their own voice mixed with noise; they were not told that the speech signal would be altered. Subjects were instructed to speak clearly, but quietly, to limit auditory feedback other than through the earphones. Subjects were also instructed to maintain normal utterance duration.

As in previous research (Houde and Jordan 1998; Purcell and Munhall 2006; Villacorta et al. 2007), we studied motor learning by providing subjects with altered auditory feedback, using a real-time acoustical transformation of the vowel sound in a CVC (consonant-vowel-consonant) utterance. The rational for using vowels sounds is that, unlike consonants, their acoustical properties can be easily modified by real-time acoustical effects processors. Vowels are distinguished by frequency peaks in their sound spectra, called "formants". Acoustical effects processors can detect and change in real time the values of these frequency peaks. For example, the vowel /ε/ in "pen" differs from the vowel /æ/ in "pan" mostly in terms of their first formant frequencies (F1). Consequently, increasing the F1 frequency value of the vowel sound /ε/ in "pen" makes it more similar acoustically to the vowel /æ/ in "pan". In the present study, subjects received auditory feedback in which the F1 frequency of the vowel in the training utterance was increased relative to the speech sounds they actually produced.

Training and transfer words and experimental conditions. Subjects were asked to repeat a single word training utterance (CVC word). Different groups of subjects were trained with different utterances. Transfer of learning was, in all cases, evaluated using the word "pen". One group of subjects was both trained and tested for transfer with the word "pen". This group served as a reference against which transfer of learning is compared. We used two different types of speech material to vary the similarity between training and testing utterances. *Experiment 1* assessed differences due to the initial consonant. *Experiment 2* examined differences due to the vowel.

In *experiment 1*, the similarity between the training and the transfer utterances was varied by using different initial consonants in the training utterance so as to assess the effects of two articulatory dimensions: voicing and place of articulation. We combined voiced and voiceless consonants with three places of articulation: bilabial, velar, and coronal. This resulted in six experimental conditions, which we tested using six different groups of subjects. In the voiceless conditions, subjects were trained either with the word "pen" (bilabial consonant, reference group), "ken" (velar consonant), or "ten" (coronal consonant). In the voiced condition, subjects were trained either with "ben", "gen", or "den". In *experiment 2*, the similarity between the training and the transfer conditions was varied by changing the vowel in the training word. We used three front vowels that differed primarily in F1 frequency: /ε/ in "pen" (reference group), /æ/ in "pan", and /I/ in "pin". Note that, in each experiment, the auditory perturbation only affected the vowel. The perturbation did not affect the identity of the consonant.

Real-time acoustical processing. As in Rochet-Capellan and Ostry (2011), an acoustical effects processor (Voice One, TC Helicon) was used to shift the signal from the microphone and play it back to subjects in real time (Fig. 1). The output of the microphone was sampled at 44,100 Hz. The speech signal was simultaneously sent both to the Voice One and to an electronic delay device. The Voice One shifted all formant frequencies, but kept the pitch unchanged. The output of the Voice One was analog low-pass filtered to preserve the pitch and the altered first formant frequency. In parallel, the signal went through a delay device that compensated for the time delay introduced by the Voice One. The delayed signal was analog high-pass filtered to preserve frequencies in the original signal higher than F1. The same cut-off frequency was used for the high-pass and low-pass filters. This frequency was determined for each subject separately, according to the value of formants in productions that were recorded and analyzed during a familiarization phase of the experiment. The outputs of the high-pass and low-pass filters were then mixed together, masking noise was added, and the resulting signal was played back to subjects. The acoustical processing took ~11 ms, which was not perceptible to the speakers. We increased the volume

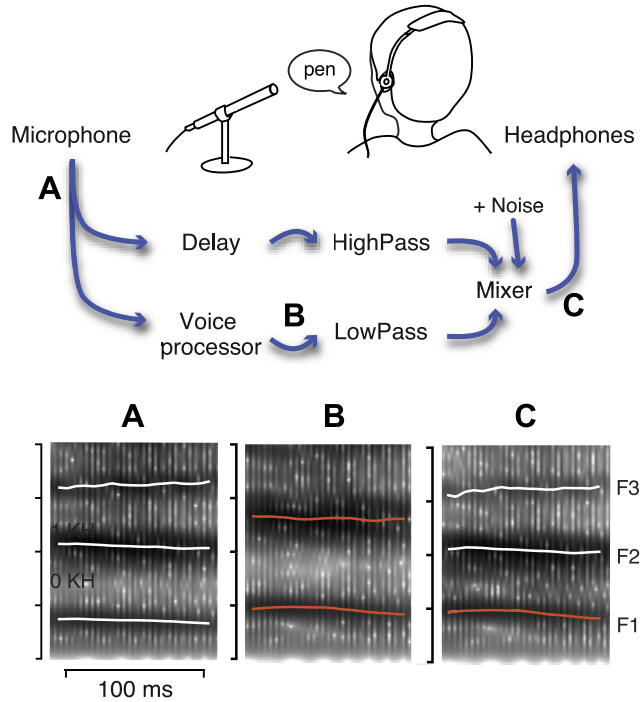

Fig. 1. Setup and real-time sound processing. Subjects read aloud words that were displayed one at a time on a monitor. The acoustical signal from the microphone was transformed in real time using a voice processor. The transformation resulted in an increase in the first formant (F1) frequency, while the pitch and higher formant frequencies were unaltered. The transformed signal was played back to the subject in real time through earphones. *A*: the original signal. *B*: the output of the voice processor. *C*: reconstituted signal that is presented to the subject. White lines are the original formant tracks. Red lines are the transformed formants.

of the signal that was played back to subjects. The volume change, along with the masking noise, helped to minimize any feedback the subject might receive other than through the earphones. The change in F1 frequency was about +25% of the original F1 value.

Setup and procedure. The experimental setup was the same as in Rochet-Capellan and Ostry (2011). Testing took place in a soundproof room. Subjects were seated at a table. They wore earphones (Stax SR001-MK2 electrostatic) and talked into a unidirectional microphone (Sennheiser, Germany). The words were displayed one at a time for 1.2 s. Two successive words were separated by 1.2 s. In a familiarization phase, subjects repeated the training and transfer words with normal feedback. The experiment consisted of 10 blocks of trials separated by 30-s pauses. The two first blocks were produced with normal feedback and contained 30 repetitions of the training word, followed by 30 repetitions of the transfer word. The auditory transformation was then introduced gradually in five discrete steps, each consisting of 10 repetitions of the training word. The frequency shift was then maintained at this maximum value for 5 blocks of 30 repetitions of the training word. After training, the auditory perturbation was turned off, and subjects were required to produce one block of 30 repetitions of the transfer word followed by one block of 30 repetitions of the training word.

Data analysis. Following data collection, the recorded signals were resampled at 10,000 Hz. We used Praat (freeware provided by Paul Boersma and David Weenink, Phonetic Sciences, University of Amsterdam, Amsterdam, The Netherlands) to detect the boundaries of vowels and then to visualize and correct these boundaries when necessary. Trials with errors of production or noise were discarded from the analyses. Formants (F1 and F2) were scored for each trial separately, using an LPC analysis on a window of 30 ms in the center of the vowel (in Praat, Burg method; Boesma and Weenink 2010). Individual utterances with F1 values beyond ±2 standard deviations of the mean in a given block were removed.

Changes in F1 frequency over the course of the experiment were expressed as a proportion of F1 values in baseline trials. A relative measure enabled us to correct for absolute differences in formant frequencies between female and male subjects that result from differences in the lengths of their vocal tracts. Learning was evaluated using the mean F1 frequency in the 30 last trials in the training phase (Learning, Fig. 2) relative to the mean F1 for that same utterance in the 30 prelearning trials. After-effects were evaluated using the same 30 baseline trials and the 30 trials for that same utterance in the no-shift postlearning phase (After-effect, Fig. 2). Transfer of learning was assessed by computing the mean F1 frequency for the transfer word under no-shift conditions in the 30 trials of the transfer phase at the end of the experiment (Transfer, Fig. 2). This value was expressed relative to the mean F1 value for the transfer utterance in the 30 prelearning trials. We used ANOVA to assess differences in frequency change in *experiments 1* and *2* separately. A single group of subjects who were trained and tested using the word "pen" was

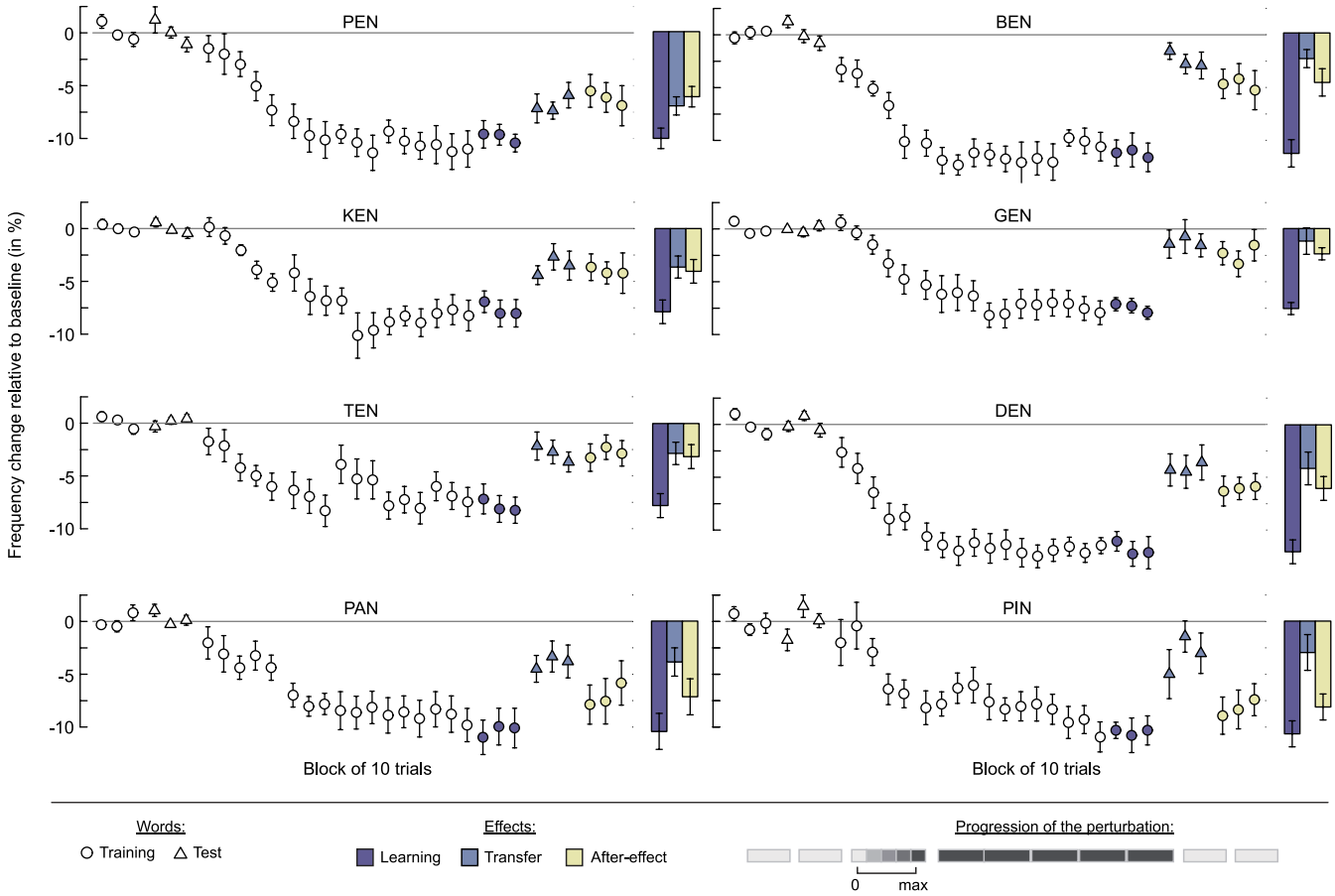

Fig. 2. Progression over time of speech motor adaptation to altered auditory feedback. In response to increases in F1 frequency in auditory feedback, subjects progressively decrease the F1 frequencies in their productions of the training word. Each panel shows the average learning curve expressed in terms of F1 frequency change relative to baseline for the different training utterances. Each symbol represents the mean value over 10 repetitions. Performance at the end of learning is shown with filled violet-colored circles. Transfer of learning (shown with blue triangles) was evaluated immediately following learning by having subjects produce the transfer utterance "pen" with unaltered auditory feedback. After-effect trials (shown with yellow circles) assess retention of learning. The bar plots at the *right* summarize the mean F1 change (±SE) at the end of learning (violet), in transfer (blue), and in after-effect (yellow) trials. The progression of the perturbation over the experiment is shown at the *bottom right*, where the gray-scale gradient represents the magnitude of F1 change. Darker colors represent greater F1 change.

included in both analyses and served as a reference group to assess transfer of learning. The inclusion of "pen" provided a continuum, starting with a zero difference between training and test material.

We evaluated the dependence of transfer of learning on the acoustic similarity between the training word and the transfer word. We used a standard measure of similarity based on the distance between vowel sounds in the training and transfer words under baseline conditions. The distance between vowels was computed on a per-subject basis using the Euclidian distance in an F1 vs. F2 space. The distance measure was then normalized, by dividing it by the F1 value in the transfer word before learning. This normalization once again accounted for possible overall differences in frequency due to female vs. male subjects. (Note that individual vowels are typically classified according to both F1 and F2 frequencies, and that formants for the same vowel vary according to the preceding consonant, due to coarticulation, Hillenbrand et al. 2001.)

It should be noted that, in studies of human arm movement, transfer of learning is generally computed on the basis of a smaller number of trials. However, in speech, there is considerable variation in formant frequencies, even for a given vowel and speaker. Accordingly, we have used a larger number of trials to deal with this variability. This variability is visible in Fig. 2, which shows the progression of learning, transfer, and after-effect magnitudes for blocks of 10 trials. To ensure that the effects reported below were not dependent on the number of trials used to compute transfer of learning, we repeated the analyses using blocks of 5, 10, and 15 trials. The results were qualitatively similar to those reported below.

Experimental condition and subject selection. As the study focused on transfer of motor learning, we only included subjects who showed clear adaptation, that is, a significant decrease in the F1 frequency in the final 30 trials of the training phase compared with the 30 trials for that same utterance under baseline conditions. As in other studies of adaptation to altered auditory feedback, not all subjects adapted. In the present study, nonadapted subjects represented ~20% of all subjects tested. We also removed 5 subjects out of the 108 who adapted. These subjects had outlier values for adaptation or transfer (1.5 times greater or less than the interquartile range of their group).

In total we retained 103 subjects who were split into groups as follows: "pen", 13 subjects (6 men); "ken", 14 subjects (6 men); "ten", 14 subjects (7 men); "ben", 13 subjects (5 men); "gen", 12 subjects (3 men); "den", 13 subjects (6 men); "pan", 12 subjects (4 men); and "pin", 12 subjects (4 men).

RESULTS

As subjects produced the training utterance, they heard F1 frequencies that were increased in real time. Figure 2 shows changes in F1 frequency over the course of the experiment for subjects in the different groups. The individual points represent the mean frequency for blocks of 10 trials. Changes in F1 frequency for the training utterance (and also the after-effect) are shown relative to baseline values for the training utterance (circles). Changes in F1 frequency for the transfer utterance are shown relative to F1 baseline values for the transfer utterance (triangles). Average values for F1 change at the end of learning; in transfer and after-effect trials are shown to the right. Examination of the figure shows a progressive reduction in F1 frequency over the course of training, which is indicative of adaptation. The F1 frequencies in the transfer phase differ for the different experimental conditions. After-effect trials show retention of motor learning.

Significant adaptation to altered auditory feedback was observed in all groups of subjects. The F1 frequencies for repetitions of the training word were significantly less at the end of training than before learning ($P < 0.0001$ in all groups). For conditions involving different consonants (Fig. 2, *top* three rows), the mean change in F1 frequency following learning ranged from 8 to 12%. For conditions involving different vowels (Fig. 2, *bottom* row), the change in F1 frequency due to learning was ~10%. In the first experiment, we evaluated differences in learning due to voiced vs. voiceless consonants (b, g, and d vs. p, k, and t) and differences due to place of articulation (labial vs. velar vs. coronal). Subjects trained with a word starting with a voiced consonant ("ben", "gen", "den") showed greater changes in the vowel production than subjects trained with a voiceless consonant ("pen", "ken", "ten") [$F(1,73) = 4.4$, $P < 0.05$]. This adaptation also depended on the place of articulation [$F(2,73) = 3.9$, $P < 0.05$], with less adjustment of the vowel observed when it was preceded by a velar consonant ("ken", "gen") than when it was preceded by a labial consonant ("pen", "ben") ($P < 0.05$). In *experiment 2*, the differences in learning for different vowels were not statistically significant: the adjustment in F1 frequency was comparable for the three words "pen", "pan", and "pin" [$F(2,34) = 0.05$, $P > 0.9$].

Transfer of learning was assessed by examining changes in F1 frequency in the production of the reference word "pen" following learning with various training words. As we observed different amount of learning in the different experimental conditions, we normalized transfer by the amount of learning. This computation was conducted on a per-subject basis. Figure 3*A* shows the mean values for transfer in the different experimental conditions. Figure 3*B* shows the relationship between transfer and the distance between vowels in the training and transfer conditions. The individual points in Fig. 3*B* give the mean value in each of the experimental conditions. The distance between training and transfer utterances is given as the Euclidian distance between vowels in F1-F2 space normalized by F1 value (see METHODS). In both panels, it is seen that transfer of learning is greatest for the reference group that was trained and tested for transfer with the same word "pen". Measures of transfer decrease in a systematic fashion with increases in the distance in sound space between training and transfer words. The details of the analyses are given below.

To test for differences in transfer in the different experimental conditions, we conducted separate ANOVAs for conditions involving consonants vs. vowels. For these analyses, we used measures of transfer that were normalized for differences in the amount of learning, as described above. When training and transfer utterances differed in terms of their initial consonant, transfer of learning was found to be greater when the training word started with an unvoiced consonant ("pen," "ken," and "ten") than when the training word started with a voiced consonant ("ben," "gen," and "den") [$F(1,73) = 6.6$, $P < 0.02$]. Transfer of learning did not differ with the place of articulation of the initial consonant [$F(2,73) = 0.31$, $P > 0.5$]. In *experiment 2*, the amount of transfer differed for training words with different vowels [$F(2,34) = 3.4$, $P < 0.05$].

We assessed the relation between transfer of learning and the acoustical similarity between the training and transfer words. Correlations were computed for the consonant (*experiment 1*) and vowel (*experiment 2*) conditions separately. When correlations were computed using condition means (as shown in Fig. 3*B*), transfer of learning for consonants was reliably correlated with the distance between training and transfer materials [$r(5) = -0.97$, $P < 0.01$]. For vowels, the correlation was

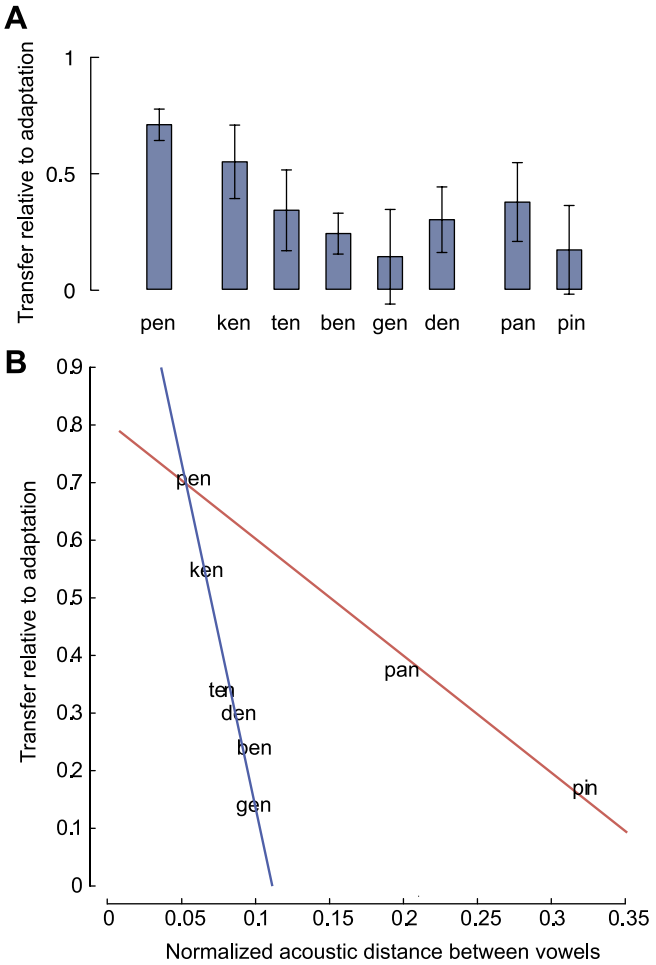

Fig. 3. Transfer of speech motor learning and its relation to similarity between training and transfer utterances. Transfer of speech motor learning decreases as a function of the distance in sound space between training and transfer utterances. *A*: mean transfer of learning (as a proportion of adaptation) to the transfer utterance "pen" for subjects who were trained with each of the different training utterances shown on the horizontal axis. The panel is divided into training words that differ in terms of the initial consonant (at the *left*) and those that differ in term of the vowel (to the *right*). Error bars show SE. *B*: transfer of learning (relative to adaptation) varies with the mean measured distance, before learning, between vowels in training and transfer utterances. Individual words represent the amount of transfer for the different training utterances. Separate regression lines are provided to characterize transfer across differences in initial consonant (blue) and transfer for different vowel conditions (red).

high, but not statistically reliable [$r(2) = -0.99$, $P = 0.08$]. When the computations were repeated using the data from individual subjects, we observed reliable correlations for both consonants [$r(78) = -0.23$, $P < 0.05$] and vowels [$r(36) = -0.42$, $P < 0.01$]. In all cases, transfer of learning decreased with the distance between training and transfer utterances.

After-effects provide a measure of the retention of motor learning. We used ANOVA to assess differences in after-effect magnitudes. We normalized after-effects on a per-subject basis by the magnitude of adaptation, to correct for differences in the amount of learning. We found no significant differences in the normalized after-effect magnitude for either consonants or vowels [$F(5,73) = 0.56$, $P > 0.7$; $F(2,34) = 0.7$, $P > 0.4$, respectively]. However, in each experimental condition, the mean after-effect magnitude was found to be reliably different than the baseline F1 value before training ($P < 0.01$ in all cases, except "ten" and "gen," where $P < 0.05$). This indicates that the learning was retained even following transfer trials in which subjects were required to produce the transfer word under normal feedback.

DISCUSSION

The results show that there is limited transfer of auditory-motor learning to untrained utterances. The generalization that was observed depended on the acoustical similarity between the two words. Gradients of generalization were observed both when training and transfer words differed in terms of the initial consonant (*experiment 1*) and in terms of the vowel (*experiment 2*). This suggests that, in the present study, speech learning is associated with individual training utterances. This work provides a new experimental model to investigate generalization in motor learning and suggests that patterns of generalization, reported previously for arm movement, are shared by the highly complex orofacial movements in speech.

Different aspects of our data point to the rather specific nature of speech motor leaning. First, speech learning has limited effects on untrained utterances. The limited generalization is consistent with previous studies that assessed transfer of speech learning after training with a mechanical load applied to the jaw (Tremblay et al. 2008). A second indication of specificity is that the effects of learning were still present for the training word after subjects repeated the transfer utterance 30 times with unaltered auditory feedback. After-effects were comparable in magnitude for the reference condition and the other experimental groups. Hence, the production of the transfer utterance did not alter the persistence of the learning effect. Finally, we have observed that transfer of learning was positively correlated with the acoustical similarity between training and transfer utterances. The after-effects, in combination with the limited amount of transfer of learning, shows that learning one word has little effect on the production of another.

In the present paper, we have assessed generalization by evaluating how learning transfers to untrained materials. The observed gradients of generalization are similar to those reported previously in the arm movement literature (Donchin et al. 2003; Gandolfo et al. 1996; Ghahramani and Wolpert 1997; Krakauer et al. 2000; Mattar and Ostry, 2007b; Thoroughman and Shadmehr 2000; Thoroughman and Taylor 2005). Other evidence for specificity (or generalization) in motor learning is seen in subjects' ability to simultaneously learn several different sensorimotor transformations, as has been observed in Rochet-Capellan and Ostry (2011) in speech or Osu et al. (2004) in arm movement. Taken together, these outcomes indicate that changes in the motor system induced by learning are primarily linked to the training experience and, hence, essentially local.

Other studies have investigated generalization of auditory-motor learning using different experimental procedures (Cai et al. 2010; Houde and Jordan 1998; Villacorta et al. 2007). In these other studies, transfer of training was tested over the course of learning by interleaving training and transfer utterances. The authors observed generalization that they interpreted as evidence that local experience induces broad modification of the speech motor system. However, in all of these previous studies, generalization to transfer utterances varied in

magnitude, and for some transfer utterances was not present at all. In fact, even with interleaving, a graded pattern of generalization can be seen in these studies as training and transfer sounds differed in acoustical similarity. Gradients in generalization provide an alternative explanation for nonhomogenous generalization in speech motor learning, consistent with the idea that learning is fundamentally instance based. Generalization of the sort seen in the present study should be expected in an instance-based system. Tuning curves for individual items in a biological system are never abrupt; changes in individual elements in a network propagate to adjacent elements as a function of proximity.

Our study confirms and extends previous findings on speech auditory-motor learning. As in previous work, the compensation for altered auditory feedback was partial. In the present data, compensation ranged between 30% for "gen" to 60% for "den". Partial compensation to altered auditory feedback may be due to the fact that speech movements have both auditory and somatosensory goals (Feng et al. 2011; Nasir and Ostry 2006; Tremblay et al. 2003). The present manipulation creates a conflict between these sources of information, which may limit adaptation. Unequal adaptation for different movements is also observed in force-field learning, for movements in different directions (Darainy et al. 2009). In the case of arm movements, differences appear to be related to limb impedance, such that, in directions where impedance is high, less adaptation is observed. Differences in adaptation in speech may well depend in part on the mechanical behavior of the articulators or the reliance on cutaneous afferent information for speech control. However, unequal adaptation may also reflect differences in the precision of implicitly defined speech targets. Note that the variation in the magnitude of adaptation in different conditions does not affect the conclusions of the present study, as differences in the amount of transfer are not due to differences in the amount of learning.

In Fig. 3*B*, it can be seen that training utterances involving different consonants are quite similar acoustically in terms of their separation in F1/F2 space (*experiment 1*), but show no more transfer than dissimilar utterances involving different vowels (*experiment 2*). This difference merits comment. The acoustical differences between training and transfer utterances in *experiments 1* and *2* have different phonetic origins. In *experiment 2*, the distance in F1/F2 space between the vowel in the training and testing utterances reflects differences in the identity of the vowel, while, in *experiment 1*, the same measure reflects the influence of the initial consonant on the vowel. One way to better evaluate the similarity between two utterances involving different consonants and the same vowel would be to characterize the associated articulatory movement. The addition of an articulatory dimension would provide another measure of similarity that could be well suited to evaluating generalization when training and test words differ in terms of their consonants.

Few studies have investigated generalization in speech motor learning. The existing literature has drawn on work on generalization in human arm movement, but adapted the techniques to create relevant experimental procedures for speech. The work on generalization in speech learning clearly benefits from the findings of the limb movement literature and, in return, helps in the understanding of fundamental principles in motor control. The comparison of different motor systems is fundamental to distinguish shared and specific properties of different motor systems. This rationale is the basis of the present work and of our laboratory's previous studies of speech motor learning (Rochet-Capellan et al. 2011; Tremblay et al. 2008). In studies of arm movement, gradual changes in amplitude or direction of movements are easily obtained by changing the position of the target in the workspace. These gradients of similarity between training and transfer movements generate gradients of generalization (Donchin et al. 2003; Gandolfo et al. 1996; Ghahramani and Wolpert 1997; Krakauer et al. 2000; Mattar and Ostry 2007b; Mattar and Ostry 2010; Thoroughman and Shadmehr 2000; Thoroughman and Taylor 2005). The present study showed equivalent gradients for speech, using a workspace in which utterances were selected in terms of their acoustical similarity. In studies of motor learning involving arm movement, broad generalization is observed when transfer movements can be interpolated from a set of training movements (Gandolfo et al. 1996; Ghahramani and Wolpert 1997; Malfait et al. 2005; Mattar and Ostry 2007a). Generalization of this kind is presumably also possible in speech when the set of training words broadly samples the articulatory workspace so as to include the articulatory movements required to realize the transfer utterance. Arm movement studies have also manipulated variables such as context to determine the factors that could influence generalization in motor learning (Cothros et al. 2009; Imamizu et al. 2007; Krakauer et al. 2006; Osu et al. 2004; Wada et al. 2003). The manipulation of higher-level factors, such as the semantic proximity between words, could be an analogous manipulation to assess contextual specificity in speech learning.

The present study has used real-time perturbation of auditory feedback to study learning and transfer. Perturbation of auditory feedback in the course of speech production also occurs in naturalistic situations. For example, in everyday life, speakers have to change their speech to compensate for ambient noise. Acoustical perturbations also result from anatomical changes of the vocal tract, in the course of child development. In natural situations, it is difficult to determine the extent of generalization of auditory-motor adaptation because individuals are exposed to many exemplars at the same time, and thus there is a constant interplay between repeated practice and novel experience. One way to understand the processes involved in speech generalization is to manipulate the overlap between training and transfer material under controlled experimental conditions. The present study is a first step toward this direction. The systematic investigation of generalization in speech motor learning may help to define the conditions for broad generalization, which is fundamental for rehabilitation protocols. It should also help in understanding the complex mapping between motor, acoustical, and linguistics units, which is still an open question in the speech literature (Smith 2006).

GRANTS

This research was supported by grants from National Institute on Deafness and Other Communications Disorders Grant DC-04669, the Natural Sciences and Engineering Research Council of Canada, and Le Fonds Québécois de la Recherche Sur la Nature et les Technologies, Québec, Canada.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

AUTHOR CONTRIBUTIONS

A.R.-C., L.R., and D.J.O. conception and design of research; A.R.-C. and L.R. performed experiments; A.R.-C., L.R., and D.J.O. analyzed data; A.R.-C., L.R., and D.J.O. interpreted results of experiments; A.R.-C. and D.J.O. prepared figures; A.R.-C. and D.J.O. drafted manuscript; A.R.-C. and D.J.O. edited and revised manuscript; A.R.-C., L.R., and D.J.O. approved final version of manuscript.

REFERENCES

Atkeson CG. Learning arm kinematics and dynamics. *Annu Rev Neurosci* 12: 157–183, 1989.

Bastian AJ. Understanding sensorimotor adaptation and learning for rehabilitation. *Curr Opin Neurol* 21: 628–633, 2008.

Boersma P, Weenink D. *Praat: doing phonetics by computer* (version 5.2.04) (Online). http://www.fon.hum.uva.nl/praat/ [2010].

Cai S, Ghosh SS, Guenther FH, Perkell JS. Adaptive auditory feedback control of the production of the formant trajectories in the Mandarin triphthong/iau/ and its patterns of generalization. *J Acoust Soc Am* 128: 2033–2048, 2010.

Cothros N, Wong J, Gribble PL. Visual cues signaling object grasp reduce interference in motor learning. *J Neurophysiol* 102: 2112–2120, 2009.

Darainy M, Mattar AAG, Ostry DJ. (2009) Effects of human arm impedance on dynamics learning, and generalization. *J Neurophysiol* 101: 3158–3168.

Donchin O, Francis JT, Shadmehr R. Quantifying generalization from trial-by-trial behavior of adaptive systems that learn with basis functions: theory and experiments in human motor control. *J Neurosci* 23: 9032–9045, 2003.

Feng Y, Gracco VL, Max L. Integration of auditory and somatosensory error signals in the neural control of speech movements. *J Neurophysiol* 106: 667–679, 2011.

Ghahramani Z, Wolpert DM. Modular decomposition in visuomotor learning. *Nature* 386: 392–395, 1997.

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

Hillenbrand J, Clark M, Nearey T. Effects of consonant environment on vowel formant patterns. *J Acoust Soc Am* 109: 748–763, 2001.

Houde JF, Jordan FM. Sensorimotor adaptation in speech production. *Science* 279: 1213–1216, 1998.

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

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

Krakauer JW, Mazzoni P, Ghazizadeh A, Ravindran R, Shadmehr R. Generalization of motor learning depends on the history of prior action. *PLoS Biol* 4: e316, 2006.

Nasir SM, Ostry DJ. Somatosensory precision in speech production. *Curr Biol* 16: 1918–1923, 2006.

Maas E, Robin DA, Austermann-Hula SN, Freedman SE, Wulf W, Ballard KJ, Schmidt RA. Principles of motor learning in treatment of motor speech disorders. *Am J Speech Lang Pathol* 17: 277–298, 2008.

Malfait N, Gribble PL, Ostry DJ. Generalization of motor learning based on multiple field exposures and local adaptation. *J Neurophysiol* 93: 3327–3338, 2005.

Mattar AAG, Ostry DJ. Neural averaging in motor learning. *J Neurophysiol* 97: 220–228, 2007a.

Mattar AAG, Ostry DJ. Modifiability of generalization in dynamics learning. *J Neurophysiol* 98: 3321–3329, 2007b.

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

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

Shadmehr R. Generalization as a behavioral window to the neural mechanisms of learning internal models. *Hum Mov Sci* 23: 543–568, 2004.

Smith A. Speech motor development: integrating muscles, movements, and linguistic units. *J Commun Disord* 39: 331–349, 2006.

Takahashi CD, Scheidt RA, Reinkensmeyer DJ. Impedance control and internal model formation when reaching in a randomly varying dynamical environment. *J Neurophysiol* 86: 1047–1051, 2001.

Thoroughman KA, Shadmehr R. Learning of action through adaptive combination of motor primitives. *Nature* 407: 742–747, 2000.

Thoroughman KA, Taylor JA. Rapid reshaping of human motor generalization. *J Neurosci* 25: 8948–8953, 2005.

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

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

Purcell DW, Munhall KG. Adaptive control of vowel formant frequency: evidence from real-time formant manipulation. *J Acoust Soc Am* 119: 2288–2297, 2006.

Rochet-Capellan A, Ostry DJ. Simultaneous acquisition of multiple auditory-motor transformations in speech. *J Neurosci* 31: 2648–2655, 2011.

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

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