J Neurophysiol 124: 1103–1109, 2020.
First published September 9, 2020; doi:10.1152/jn.00457.2020.

RESEARCH ARTICLE | *Control of Movement*

Contribution of sensory memory to speech motor learning

Takayuki Ito,[1,2] Jiachuan Bai,[1] and David J. Ostry[2,3]

[1]*Laboratoire de Recherche Grenoble, Images, Parole, Signal, Automatique, Grenoble Institute of Technology, Université Grenoble Alpes, Centre National de la Recherche Scientifique, Grenoble, France;* [2]*Haskins Laboratories, New Haven, Connecticut; and* [3]*McGill University, Montréal, Québec, Canada*

Submitted 30 July 2020; accepted in final form 3 September 2020

Ito T, Bai J, Ostry DJ. Contribution of sensory memory to speech motor learning. *J Neurophysiol* 124: 1103–1109, 2020. First published September 9, 2020; doi:10.1152/jn.00457.2020.—Speech learning requires precise motor control, but it likewise requires transient storage of information to enable the adjustment of upcoming movements based on the success or failure of previous attempts. The contribution of somatic sensory memory for limb position has been documented in work on arm movement; however, in speech, the sensory support for speech production comes from both somatosensory and auditory inputs, and accordingly sensory memory for either or both of sounds and somatic inputs might contribute to learning. In the present study, adaptation to altered auditory feedback was used as an experimental model of speech motor learning. Participants also underwent tests of both auditory and somatic sensory memory. We found that although auditory memory for speech sounds is better than somatic memory for speechlike facial skin deformations, somatic sensory memory predicts adaptation, whereas auditory sensory memory does not. Thus even though speech relies substantially on auditory inputs and in the present manipulation adaptation requires the minimization of auditory error, it is somatic inputs that provide the memory support for learning.

NEW & NOTEWORTHY In speech production, almost everyone achieves an exceptionally high level of proficiency. This is remarkable because speech involves some of the smallest and most carefully timed movements of which we are capable. The present paper demonstrates that sensory memory contributes to speech motor learning. Moreover, we report the surprising result that somatic sensory memory predicts speech motor learning, whereas auditory memory does not.

altered auditory feedback; auditory memory; somatosensory memory; speech motor adaptation

INTRODUCTION

The need to briefly retain information about prior movements and states is central to motor skill acquisition as it permits corrections and adjustments from one movement to the next, over the course of learning. This need is particularly clear in speech motor learning, which occurs without visual guidance and is thus potentially reliant on both auditory and somatic sensory memory. The involvement of sensory memory in motor learning has been reported in the context of upper extremity movements (Anguera et al. 2010; Bo et al. 2011; Bo and Seidler 2009; Christou et al. 2016; Sidarta et al. 2018). However, little at all is known about the contribution of sensory memory storage in relation to speech motor learning (Guo et al. 2017).

The transient retention of information, on the order of seconds or less, is required during motor learning to enable us to repeat successful movements and avoid repeating errors. Retention on this time scale has been reported for each of auditory, visual, and somatic sensory memory (Bliss et al. 1966; Crowder 1982; Gilson and Baddeley 1969; Sperling 1960). This brief retention of information can be distinguished from the short-term retention of learning, which lasts on the order of hours and precedes more durable storage of consolidated motor memory (Brashers-Krug et al. 1996). A number of studies of upper-limb movement have focused on this transient storage and its relationship to learning. Visuospatial memory capacity was found to correlate with the rate of sequence learning and visuomotor adaptation (Bo et al. 2011; Bo and Seidler 2009). In adaptation learning, this relationship may be dependent on explicit strategies and cognitive factors since it was found to be present when perturbations were introduced abruptly but not with gradually introduced perturbations that reduce awareness of the perturbation and the likelihood of using explicit strategies (Christou et al. 2016). In work on reinforcement learning, in which participants' memories were assessed for their own previously produced movements, it was found that retention was transient and effectively limited to about the single most recent movement, but nevertheless individual differences in memory capacity were correlated with the overall magnitude of learning (Sidarta et al. 2018). Taken together, these studies indicate that better sensory memory is associated with better learning and are consistent with the idea the transient storage of sensory information contributes directly to motor learning.

Transient retention of either one or both of auditory and somatosensory information is necessary for speech motor adaptation in which learning requires trial-to-trial adjustments to motor commands to deal with error. Auditory and somatic sensory memory might both contribute to learning. Somatic memory may help speakers tolerate somatic errors, which is necessary for adaptation to altered auditory feedback. Auditory memory may contribute to the repetition of utterances that sounded right on the previous trial. In the present study, we tested the hypothesis that sensory memory supports speech motor learning. Better memory performance should predict better learning, and accordingly the sensory dimensions on which a relationship is observed indicate the sensory basis of learning. We assessed adaptation to altered auditory feedback in speech and in the same subjects obtained

---

Correspondence: T. Ito (takayuki.ito@gipsa-lab.grenoble-inp.fr).

estimates of sensory memory for speech sounds and for orofacial somatic inputs. As has been reported previously, we (Lametti et al. 2012) found that speakers showed varying degrees of adaptation to altered auditory feedback. We also found that auditory sensory memory was better than somatic sensory memory. However, both kinds of memory were limited in capacity. The key finding was that differences in somatic sensory memory predicted learning, whereas auditory sensory memory did not. This is consistent with the idea that even though speech communication is auditory in nature, somatic inputs are central to speech motor learning.

MATERIALS AND METHODS

Subjects and procedure. The protocol for this experiment was approved by the local ethical committee of the Université Grenoble Alpes [Comité d'Ethique pour la Recherche, Grenoble Alpes (CERGA: Avis-2018-12-11-4)]. Twenty-one native speakers of French (nine women, ages 18–35 yr) were tested. Subjects reported no impairment of hearing or speech and signed the corresponding consent form.

We carried out three separate tests to examine performance in speech motor adaptation and sensory memory. Separate tests of auditory and somatosensory memory were conducted. We examined the correlation between speech motor adaptation and sensory memory performance.

Task utterances. We focused on the vowels /é/-/ɛ/-/a/, which are acoustically and articulatorily neighbors. In the acoustic domain, vowel sounds are characterized by vocal tract resonances known as formants that are peaks in the spectrum of speech sounds. The lowest two formants (F1 and F2) are most important in distinguishing different vowels. For the vowels in the present study, as we pass from /é/ to /ɛ/ to /a/, the value of F1 is increased, and F2 is decreased in a nearly linear fashion. Articulatorily, those acoustic changes are achieved mainly by changing the vertical position of the jaw and tongue. The production of /é/ requires the highest position of the jaw and tongue (the narrowest vocal tract constriction). The articulator positions gradually change in a downward direction from /é/ to /ɛ/ and /a/. The experimental manipulations focused on the middle sound of /ɛ/ since this sound can be altered acoustically and articulatorily in both directions (to /é/ or to /a/) in speech production. In the speech motor adaptation task, we used the French utterance /tɛ/, "taie" (pillow cover in English), for which the vowel sound is similar to that used previous studies (Purcell and Munhall 2006; Rochet-Capellan and Ostry 2011). In the auditory memory task, we tested four variants of the /tɛ/ sound on a synthesized continuum between /é/ and /a/ as described below. In the somatosensory memory task, we tested four different amplitudes of facial skin deformation along a vertical axis. We chose this direction as it corresponds to the vertical articulatory dimension that is dominant across vowels used here, as noted above. As shown previously, this manipulation provides articulatory information related to the direction of skin deformation (Ito et al. 2009; Ito and Ostry 2010, 2012).

Speech motor adaptation task. We used altered auditory feedback to assess adaptation to speech sounds altered in real time (Fig. 1*A*). Subjects were instructed to produce the utterance /tɛ/ in response to a visual cue. An intertrial interval was varied between 1 and 2 s to prevent anticipation of the start of each trial.

The altered auditory feedback manipulation used the Audapter software developed by Cai et al. (2011). Subjects wore a headset (Audio-Technica BPHS1) and spoke into a microphone on the headset. Audio signals were digitally sampled at 44,100 Hz using a universal serial bus audio interface (Steinberg UR22mkII) and then downsampled at 11,025 Hz to reduce processing time by the software. The software extracts the first and second formants using linear predictive coding. The sounds produced by the subject and those played back to subjects were both recorded, as were the extracted formants. The extracted formats were used in the analysis described below. In the altered feedback manipulation, we systematically decreased the frequency of the first formant alone so as to change the sound played back to the subject to something in the direction of /té/ (see an example of formant change in Fig. 1*A*). Our expectation was that adaptation would result in utterances that sound more like /ta/ and have associated increases in the formant frequency of the produced sound. The maximum shift was set on a per-subject basis to 15% relative to their first formant frequency. The sound volume of the auditory feedback was increased to a fixed value that was maintained over the adaptation trials and served to minimize the contribution to the signal at the cochlea of unaltered airborne and bone-conducted speech. A 70-dB masking noise was also added to the signal at the headphones.

The initial 30 trials provided a baseline phase with no auditory feedback alteration. Over the next 50 trials, the first formant frequency in the auditory feedback signal was gradually decreased (ramp phase). The maximum frequency change was then maintained for the following 50 trials (hold phase). The auditory feedback signal was abruptly returned to normal in the last 20 trials (aftereffect phase).

Sensory memory test. We carried out memory tests of audition and somatosensation separately (Fig. 1*B*). Both tests involved the same procedure. The subject's task was to identify whether a test stimulus was included in a previously presented 2-item memory set. The test item and memory set stimuli were chosen from a fixed set of 4 sensory

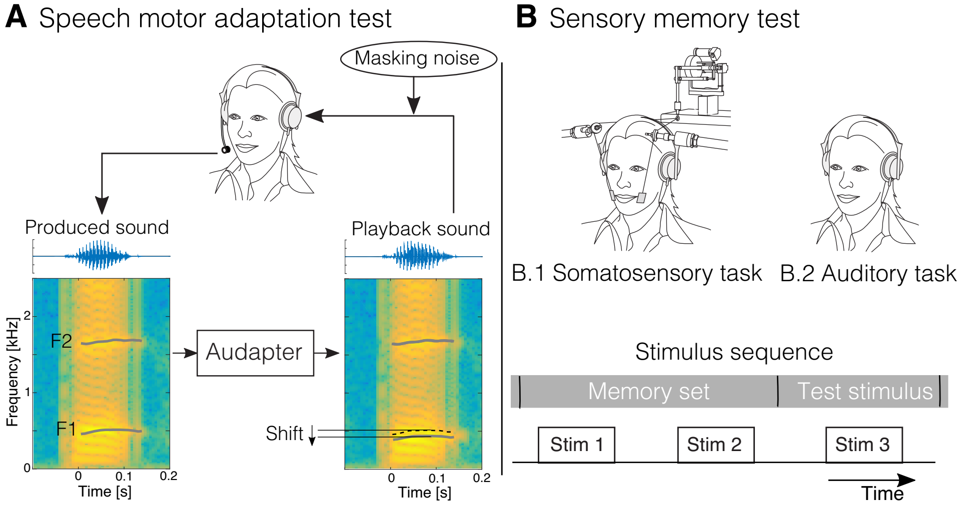

Fig. 1. Experimental setup and procedure for speech motor adaptation test (*A*) and sensory memory test (*B*). *A*: schematic view of altered auditory system. Time responses and spectrograms represent an example of produced (left) and playback sounds (right). The thick lines in the spectrogram represent the first formant (F1) and second formant (F2). The dotted line in the spectrogram of playback sound represents the original F1 shown in produced sound. *B*: experimental setup for somatosensory task (*B.1*) and auditory task (*B.2*). The bottom part represents stimulus (Stim) sequence of memory set and test stimulus in 1 trial.

stimuli. On each trial, 2 of the 4 were chosen as the memory set, and 1 of the 4 was chosen as the test stimulus. Stimulus sequence is represented at the bottom of Fig. 1*B*. All possible combinations of the 4 stimuli (48 combinations) were tested in random order and repeated 4 times each. One hundred ninety-two trials were recorded in total. The interstimulus interval was set at 500 ms throughout.

For the auditory memory tests, we chose 4 stimuli from a 19-step synthesized continuum extending from /té/ through /tɛ/ to /ta/. The continuum was generated by dividing the first and second formant frequencies between /té–tɛ/ and /tɛ–ta/ in equal parts. The continuum was based on recordings of /té/, /tɛ/, and /ta/ spoken by a native male French speaker. The stimuli were numbered from 1 to 19 (1 = té, 10 = tɛ, and 19 = ta). The average difference between any two sequential synthesized utterances was 16.5 Hz for the first formant (3.3% change relative to F1 in /ɛ/: 506.5 Hz) and 35 Hz for the second formant (2.0% relative to F2: 1,754.6 Hz). Since the speech utterance used in the adaptation task was /tɛ/, we selected as auditory memory stimuli a set of four auditory stimuli near to /ɛ/ (nos. 7, 9, 11, and 13). Previous work has shown that discrimination thresholds are 14 Hz for F1 and 31 Hz for F2 (Kewley-Port and Watson 1994). The difference between adjacent stimuli in the memory test (33 Hz in F1 and 70 Hz in F2) was well above the threshold for vowel discrimination. In a separate control test, we verified that the memory test stimuli were readily discriminable (see RESULTS).

For the somatosensory memory test, we applied facial skin stretch using a robotic device (Phantom 1.0; SensAble Technologies; see Fig. 1*B.1* for an example view of setup). The details of somatosensory stimulation setup are described in a previous study (Ito et al. 2009). Briefly, plastic tabs (2 × 3 cm) were attached on the skin lateral to the oral angle on each side of the face. These tabs were connected to the robotic device through thin wires. The wires were supported by wire supports to avoid contact with the facial skin. The skin was stretched when the robotic device applied force to the wires. The temporal profile of the applied force was a single cycle of a 3-Hz sinusoid. Since the vowels in the auditory memory test can be distinguished by their vertical elevation (during production), we applied the skin stretch in an upward direction and varied the stretch magnitude. We chose four stimuli with different peak amplitudes (0.55, 0.85, 1.15, and 1.45 N). The 0.3-N difference between adjacent skin stretch stimuli was selected to be readily discriminable, based on a previous finding in which sequential skin stretch stimuli separated by 0.2 N produced a 90% discrimination rate (Ito and Ostry 2012).

Experimental procedure. We carried out the speech motor adaptation task at the beginning of the test session followed in order by the somatosensory memory test and the auditory memory task. The fact that auditory sensory memory was found to be better overall than somatic sensory memory (see below) suggests there was no fundamental decrement in memory processing that occurred as result of the testing order (see RESULTS for additional control study). The same auditory setup was used both for the speech motor adaptation and auditory memory task, although masking noise was not used during the memory test. In the adaptation task, subjects were instructed to repeat aloud the French word taie. In the memory tests, following stimulation, subjects were instructed to indicate whether the test item was in the memory set.

Data analysis. Performance in altered auditory feedback training was evaluated using changes in the first formant frequency of the produced vowel over the course of the training. The formants in each trial were obtained by taking an average over a time window of 31.9 ms, which was centered on the root-mean-square peak of the vowel. We focused on the first formant since the frequency shift was applied in it. The obtained formants were normalized by dividing by a baseline measure for which we used the mean value over the last 10 trials of unaltered feedback (21st to 30th trials). The amount of adaptation was quantified by averaging the normalized amplitude at the end of the hold phase (average of the last 10 trials) and at the beginning of the aftereffect phase (average of the first 10 trials).

We assessed adaptation on a per-subject basis by conducting 1-tailed t tests with $\alpha = 0.05$ on the mean of the trials in the baseline phase and those at the end of the adaptation phase. As in previous studies (Lametti et al. 2012), we found that ~30% of the subjects did not adapt to the auditory feedback manipulation.

We quantified performance in the sensory memory task on a per-subject basis using the sensitivity index, d' (Macmillan and Creelman 2004), which provides a composite performance measure that combines the proportion of times that a subject correctly responds that the test stimulus was in memory set (hit rate) and the proportion of cases in which a nonmemory set item is incorrectly labeled as being part of the memory set (false alarm rate).

Correlation analyses were carried out with Pearson product moment correlation coefficient to examine the relationship between the amount of the adaptation and performance in each of the auditory and somatosensory memory tests.

RESULTS

We will consider first speech motor learning performance in the altered auditory feedback manipulation. To clarify the patterns of adaptive behavior, we divided subjects into 2 groups, adapted and nonadapted, and averaged within each group. Based on a per-subject analysis using t tests, 14 of 21 subjects (66%) were categorized as adapting with performance at the end of training reliably different from baseline. This percentage is consistent with previous studies of adaptation in speech motor learning in which 60–80% of subjects adapt (Lametti et al. 2012). Figure 2*A* shows the averaged change in the produced F1 frequency over the course of training. The amplitude of the effect at the end of the training and aftereffect phases are summarized in Fig. 2*B*. As in previous studies, the averaged F1 changes in adapted group showed a gradual change over the ramp phase, and the maximum change was maintained during the hold phase. In contrast, the nonadapted group showed no change in F1 over the course of the training trials. F1 at the end of the training was significantly different between these two groups by two-tailed Welch t test [$t(18.9) = 6.60$, $P < 0.001$]. The change in F1 was reliably different from the baseline value of 1 in the adapted group [$t(13) = 8.91$, $P < 0.001$] but not in the nonadapted group [$t(6) = 1.10$, $P = 0.32$]. In aftereffect phase, although the mean F1 in adapted group slightly returned toward the baseline value, the difference between the two groups remained reliable [$t(14.6) = 3.38$, $P < 0.01$], and the change in F1 was greater than the baseline value of 1 in the adapted group [$t(13) = 4.67$, $P < 0.001$] but not in the nonadapted group [$t(6) = 0.422$, $P = 0.69$]. These tests were also conducted across all subjects together. The change in F1 was reliably different from the baseline value of 1 following adaptation [$t(20) = 5.76$, $P < 0.001$] and remained different from baseline in the aftereffect phase [$t(20) = 3.18$, $P < 0.01$].

Sensory memory was evaluated with a commonly used measure of sensitivity, d'. Mean d' across subjects for the auditory memory task was 1.15 ± 0.07 (SE), and d' for the somatosensory memory task was 0.91 ± 0.06 (SE; Fig. 3*B*). A paired t test showed significantly greater memory performance in the auditory memory task [$t(20) = 2.64$, $P < 0.02$]. Thus somatosensory memory performance was worse than auditory memory performance in the current comparison.

In the correlation analyses, we found a reliable positive correlation between somatosensory memory and speech motor adaptation ($r = 0.52$, $P < 0.02$). Participants with better somatic memory showed greater adaptation. The correlation between

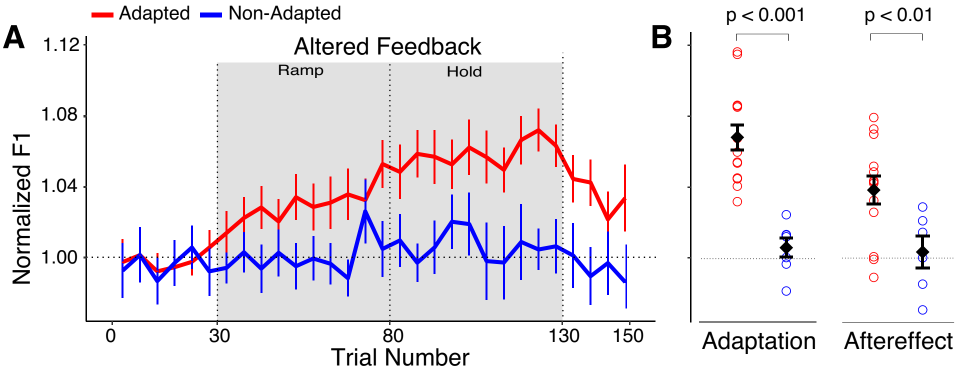

Fig. 2. Changes in first formant (F1) frequency of produced sounds in a speech motor adaptation task using altered auditory feedback. *A*: changes in the first formant over the course of trials. The error bars represent standard errors across the subjects. *B*: mean normalized frequency of the first formant at the end of the adaptation trials and in the aftereffect phase. Adapted subjects are shown in red, and nonadapted subjects in blue.

auditory memory and adaptation was not reliable ($r = 0.08$, $P > 0.7$). There was no correlation between auditory and somatosensory memory performance ($r = 0.016$, $P > 0.9$). Figure 3*A* shows scatter plots, regression lines, and 95% confidence intervals for all comparisons. The results suggest a contribution of somatosensory working memory to speech motor adaptation.

Although the difference in formant values between any two adjacent auditory stimuli that were used in the memory test was well above the discrimination threshold reported in the previous studies (Kewley-Port and Watson 1994), the lack of a correlation between auditory memory and learning may be due to a problem in discriminating the stimuli that we used for the memory test. To assess this possibility, we ran a control study in which we examined whether our stimuli are discriminable by testing discrimination performance using an ABX procedure. The subjects tested using this procedure were asked to indicate whether *stimulus X* was the same as either *stimulus A* or *stimulus B*. We tested all possible pairs of the 4 stimuli that we used in the memory test. The correct performance rate for adjacent stimulus pairs averaged across 15 subjects was 0.75 ± 0.019 (SE). The average proportion correct was 0.96 ± 0.010 (SE) for all other pairs. The proportion of correct responses was clearly above chance, although the rate for correct response with the adjacent stimulus pair was significantly smaller than the rate obtained in all other pairs [$t(14) = 13.2$, $P < 0.001$]. We thus concluded that the auditory stimulus pairs used in the memory test were discriminable, and this was not the cause of the lack of correlation between auditory memory and learning.

As a further control, to rule out the possibility that the order of perceptual testing affected measures of auditory sensory memory, we repeated the auditory memory test on its own. Ten native French speakers were tested. The obtained d' was

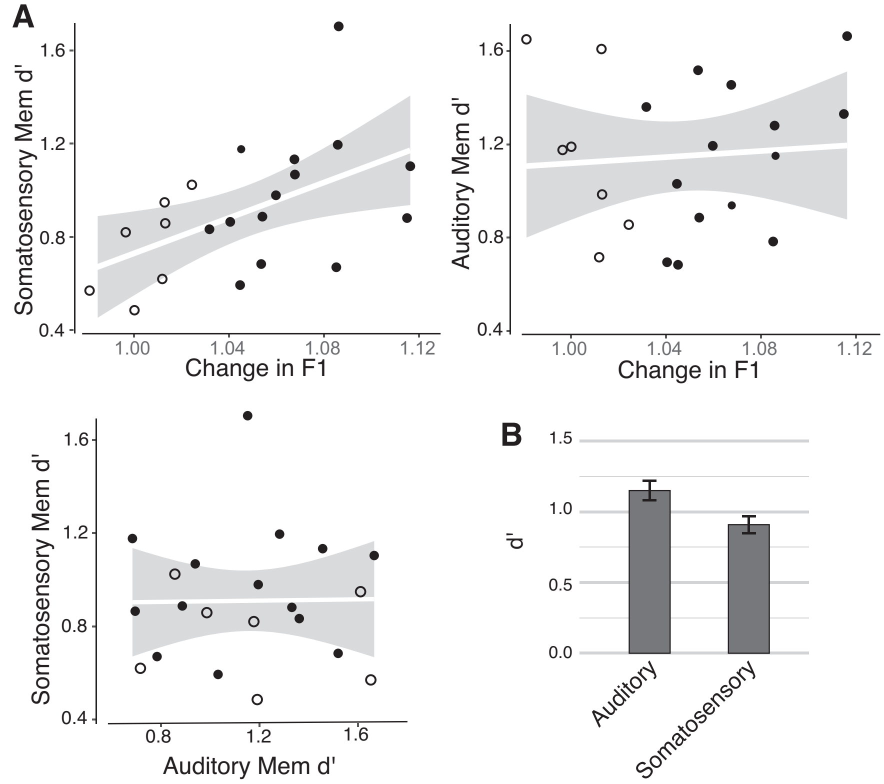

Fig. 3. *A*: scatter plots with regression lines in correlation analyses showing the relationship between learning (change in first formant, F1) and sensory memory (Mem) as well as between sensory memory measures. The shaded area represents the 95% confidence interval in the correlation analyses. The open circles give values for nonadapted subjects. Closed circles show adapted subjects. *B*: sensory memory performance, sensitivity index (d') in auditory and somatosensory memory tasks. Error bars represent the standard error across the subjects.

1.20 ± 0.12 (SE). This was not significantly different from the value of 1.15 ± 0.07 (SE) in the main test described above [t (15.3) = 0.37, $P > 0.7$]. This suggests that the auditory memory score was not biased by the order in which the tests were conducted.

DISCUSSION

The present study tested for the involvement of sensory memory in speech motor learning using adaptation to altered auditory feedback as an experimental model of learning. We carried out within-subject tests in which we obtained measures of adaptation along with both somatic and auditory sensory memory with the goal of assessing their relative contributions to learning. The stimuli used in the memory task were discriminable, but the task was intentionally difficult, to simulate the often subtle differences between correct and incorrect movements and sounds that normally occur during speech motor learning. We found that overall, subjects' auditory sensory memory was better than their somatic sensory memory. Sensory memory scores were uncorrelated. As has been reported previously, the magnitude of adaptation to altered auditory feedback varied substantially (Lametti et al. 2012). Somatic memory scores predicted the amount of adaptation, whereas auditory memory scores did not.

The results are consistent with the substantial involvement of the somatosensory system in speech motor learning, and this is a main finding of the present study. The findings complement an accumulating body of evidence that points to the contribution of the somatosensory system to both speech learning and its control (Darainy et al. 2019; Feng et al. 2011; Ito et al. 2016; Ito and Ostry 2010; Jones and Munhall 2003; Tourville et al. 2008; Tremblay et al. 2003, 2008). It has been shown previously that perturbations to the somatosensory system during speech movements result in online adjustments to movement (Abbs et al. 1984; Folkins and Abbs 1975; Honda et al. 2002; Ito et al. 2020). There is a somatic contribution to learning that is seen in response to predictable somatosensory errors during adaptation and occurs even in the absence of auditory feedback (Baum and McFarland 1997; Brunner et al. 2011; Feng et al. 2011; Hamlet et al. 1976; Ito and Ostry 2010; Jones and Munhall 2003; Nasir and Ostry 2006, 2008; Savariaux et al. 1995; Tremblay et al. 2003, 2008). Individuals differ in their reliance on auditory versus somatosensory information for adaptation (Lametti et al. 2012). There are also speech-learning-related changes to somatosensory areas of the brain (Darainy et al. 2019; Ito et al. 2016). The dependence of speech, which is fundamentally acoustic, on somatic information presumably arises over the course of learning. Individual movements even in the absence of perturbations result in correlated auditory and somatosensory feedback. This presumably contributes to the acquisition and retention of what can be termed somatic targets or goals or expectations for movements in addition to expectations regarding the associated sounds. This would account for changes to both somatic areas of the brain in the course of speech motor learning and to movement during speech as a result of somatosensory perturbations.

The involvement demonstrated here of somatic sensory memory in speech motor learning does not rule out the potential participation of auditory sensory memory in other speech-learning-related tasks. The adaptation task used here may be relatively insensitive to the contribution of auditory memory. Adaptation to altered auditory feedback requires a speaker to tolerate somatic error to make the stimuli sound right. Accordingly, to sound right, somatic error is generally large when auditory error is small. Thus it could be expected that speakers with better somatic memory are better able to tolerate somatic error to produce movements that minimize the error in the target sounds. An experimental model such as second-language learning may be better able to equate somatic and auditory discrepancies that occur over the course of learning.

A number of previous studies have assessed the relation between sensory memory and aspects of movement production. In the context of speech, Ranasinghe et al. (2019) examined the relationship between reflex compensation for pitch perturbation and measures working memory performance in patients with Alzheimer's disease and found that deficits in pitch compensation were related to those in memory performance. More information on this relation comes from work on upper-limb movement, in which differences in sensory memory between individuals have shown to be related to individual differences in learning in adaptation, sequence learning, and reinforcement learning tasks (Bo et al. 2011; Bo and Seidler 2009). In limb-movement studies, the extent to which the memory support for learning engages cognitive mechanisms is unknown. One previous study showed a possible dependence of the relationship between memory and motor learning situations in which explicit strategies might play a role (Bo et al. 2011; Bo and Seidler 2009).

In the present study, the auditory shift was introduced gradually to minimize subject awareness of the perturbation with the goal of assessing implicit learning processes. Abruptly introduced auditory perturbations have also been used in speech adaptation studies. These result in clearly detectable acoustic changes and presumably tap into any cognitive strategies (explicit learning) as well as implicit learning processes. However, there are few differences in the magnitude of adaptation under these conditions, compared with adaptation observed to gradually introduced shifts (Keough et al. 2013; MacDonald et al. 2010). Moreover, even when subjects are instructed to ignore altered auditory feedback, they are unable to do so (Keough et al. 2013; Munhall et al. 2009). This suggests abrupt perturbations are unable to engage learning processes other than those involved in gradually introduced loads, and, in turn, this suggests that the relationship observed here between learning and sensory memory likely taps into implicit elements of the learning process.

The present paper uses measures of overall sensory memory capacity rather than memory as it relates to trial-to-trial learning. This same approach has been taken in studies that have examined the relationship between upper-extremity movement and human motor learning (Anguera et al. 2010; Bo et al. 2011; Bo and Seidler 2009; Christou et al. 2016). We know of only one study (Sidarta et al. 2018) in which the relationship between self-produced movements and learning is examined. This study tested memory by periodically doing it over the course of learning by using a robot to passively play back candidate movements and then asking subjects to judge whether the displacement produced by the robot was to the right or the left of a movement they had just produced. As in other reports, memory correlated positively with learning performance. This same approach could be applied in the context of speech motor learning, at least to test for memory for produced sounds, and would be a worthwhile undertaking in future studies. However, directly comparable tests for somatic sensory memory in speech would be difficult if not impossible because of the challenge of transducing the original movements (in the case of the facial skin or tongue) and the inability to mechanically displace structures such as the jaw. The present scientific question required measures of both auditory and somatic memory, and accordingly we chose composite measures rather than trial-by-trial assessments.

It is presently unknown whether the sensory memory support for learning is within areas of the brain that are engaged in the learning process itself. The idea that local, region-specific memory processes contribute to learning is consistent with work by Romo and de Lafuente (2013), who have shown that somatic memory and decision making occur over a wide range of areas with parietal and frontal cortex, which include second somatosensory cortex, ventral premotor cortex, supplementary motor area, and ventrolateral prefrontal cortex. However, the large literature on hippocampal involvement in memory and on dorsolateral prefrontal cortex in a variety of tasks involving retention suggests a dissociation may be possible. Whether local sensory memory or a separate sensory store enables speech learning is currently unknown.

The sensory memory tasks reported in this study were both run after adaptation to avoid possible effects on sensory memory testing on learning. The somatic memory task was run first for methodological reasons to provide subjects with a break during the setup of the skin-stretch apparatus before the memory tests. Although the order effect is a limitation and could have affected the correlation, the fact that auditory memory is better than somatosensory memory suggests that the order of testing has not degraded auditory memory. As a control, we repeated auditory memory tests on their own. The obtained auditory sensory memory magnitude was similar to that in the main study, which suggests that the testing order did not introduce bias.

A test was conducted to evaluate the discriminability of the auditory stimuli that were used in the sensory memory test. The goal was to rule out the possibility that auditory sensory memory failed to correlate with adaptation because, on some dimension, the stimuli were not discriminable (even though auditory memory was better). A control study was not run for the somatosensory condition because it was found be correlated with learning (so there was no failure to explain). Moreover, in other work, we (Ito and Ostry 2012) had already reported tests showing that skin-stretch magnitudes even less than those used in the present study were readily discriminable.

GRANTS

This work was supported by the National Institute on Deafness and Other Communication Disorders Grant R01-DC017439, the European Research Council under the European Commission's Seventh Framework Program (FP7/2007-2013 Grant Agreement no. 339152), and NeuroCoG Initiatives d'Excellence Université Grenoble Alpes in the framework of the "Investissements d'Avenir" Program (ANR-15-IDEX-02).

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the authors.

AUTHOR CONTRIBUTIONS

T.I. and D.J.O. conceived and designed research; T.I. and J.B. performed experiments; T.I. and J.B. analyzed data; T.I., J.B., and D.J.O. interpreted results of experiments; T.I. and J.B. prepared figures; T.I. and D.J.O. drafted manuscript; T.I. and D.J.O. edited and revised manuscript; T.I. and D.J.O. approved final version of manuscript.

REFERENCES

Abbs JH, **Gracco VL**, **Cole KJ.** Control of multimovement coordination: sensorimotor mechanisms in speech motor programming. *J Mot Behav* 16: 195–232, 1984. doi:10.1080/00222895.1984.10735318.

Anguera JA, **Reuter-Lorenz PA**, **Willingham DT**, **Seidler RD.** Contributions of spatial working memory to visuomotor learning. *J Cogn Neurosci* 22: 1917–1930, 2010. doi:10.1162/jocn.2009.21351.

Baum SR, **McFarland DH.** The development of speech adaptation to an artificial palate. *J Acoust Soc Am* 102: 2353–2359, 1997. doi:10.1121/1.419619.

Bliss JC, **Crane HD**, **Mansfield PK**, **Townsend JT.** Information available in brief tactile presentations. *Percept Psychophys* 1: 273–283, 1966. doi:10.3758/BF03207391.

Bo J, **Jennett S**, **Seidler RD.** Working memory capacity correlates with implicit serial reaction time task performance. *Exp Brain Res* 214: 73–81, 2011. doi:10.1007/s00221-011-2807-8.

Bo J, **Seidler RD.** Visuospatial working memory capacity predicts the organization of acquired explicit motor sequences. *J Neurophysiol* 101: 3116–3125, 2009. doi:10.1152/jn.00006.2009.

Brashers-Krug T, **Shadmehr R**, **Bizzi E.** Consolidation in human motor memory. *Nature* 382: 252–255, 1996. doi:10.1038/382252a0.

Brunner J, **Ghosh S**, **Hoole P**, **Matthies M**, **Tiede M**, **Perkell J.** The influence of auditory acuity on acoustic variability and the use of motor equivalence during adaptation to a perturbation. *J Speech Lang Hear Res* 54: 727–739, 2011. doi:10.1044/1092-4388(2010/09-0256).

Cai S, **Ghosh SS**, **Guenther FH**, **Perkell JS.** Focal manipulations of formant trajectories reveal a role of auditory feedback in the online control of both within-syllable and between-syllable speech timing. *J Neurosci* 31: 16483–16490, 2011. doi:10.1523/JNEUROSCI.3653-11.2011.

Christou AI, **Miall RC**, **McNab F**, **Galea JM.** Individual differences in explicit and implicit visuomotor learning and working memory capacity. *Sci Rep* 6: 36633, 2016. doi:10.1038/srep36633.

Crowder RG. A common basis for auditory sensory storage in perception and immediate memory. *Percept Psychophys* 31: 477–483, 1982. doi:10.3758/BF03204857.

Darainy M, **Vahdat S**, **Ostry DJ.** Neural basis of sensorimotor plasticity in speech motor adaptation. *Cereb Cortex* 29: 2876–2889, 2019. doi:10.1093/cercor/bhy153.

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. doi:10.1152/jn.00638.2010.

Folkins JW, **Abbs JH.** Lip and jaw motor control during speech: responses to resistive loading of the jaw. *J Speech Hear Res* 18: 207–220, 1975. doi:10.1044/jshr.1801.207.

Gilson EQ, **Baddeley AD.** Tactile short-term memory. *Q J Exp Psychol* 21: 180–184, 1969. doi:10.1080/14640746908400211.

Guo Z, **Wu X**, **Li W**, **Jones JA**, **Yan N**, **Sheft S**, **Liu P**, **Liu H.** Top-down modulation of auditory-motor integration during speech production: the role of working memory. *J Neurosci* 37: 10323–10333, 2017. doi:10.1523/JNEUROSCI.1329-17.2017.

Hamlet SL, **Geoffrey VC**, **Bartlett DM.** Effect of a dental prosthesis on speaker-specific characteristics of voice. *J Speech Hear Res* 19: 639–650, 1976. doi:10.1044/jshr.1904.639.

Honda M, **Fujino A**, **Kaburagi T.** Compensatory responses of articulators to unexpected perturbation of the palate shape. *J Phon* 30: 281–302, 2002. doi:10.1006/jpho.2002.0172.

Ito T, **Coppola JH**, **Ostry DJ.** Speech motor learning changes the neural response to both auditory and somatosensory signals. *Sci Rep* 6: 25926, 2016. doi:10.1038/srep25926.

Ito T, **Ostry DJ.** Somatosensory contribution to motor learning due to facial skin deformation. *J Neurophysiol* 104: 1230–1238, 2010. doi:10.1152/jn.00199.2010.

Ito T, **Ostry DJ.** Speech sounds alter facial skin sensation. *J Neurophysiol* 107: 442–447, 2012. doi:10.1152/jn.00029.2011.

Ito T, **Szabados A**, **Caillet JL**, **Perrier P.** Quick compensatory mechanisms for tongue posture stabilization during speech production. *J Neurophysiol* 123: 2491–2503, 2020. doi:10.1152/jn.00756.2019.

Ito T, **Tiede M**, **Ostry DJ.** Somatosensory function in speech perception. *Proc Natl Acad Sci USA* 106: 1245–1248, 2009. doi:10.1073/pnas.0810063106.

Jones JA, **Munhall KG.** Learning to produce speech with an altered vocal tract: the role of auditory feedback. *J Acoust Soc Am* 113: 532–543, 2003. doi:10.1121/1.1529670.

Keough D, **Hawco C**, **Jones JA.** Auditory-motor adaptation to frequency-altered auditory feedback occurs when participants ignore feedback. *BMC Neurosci* 14: 25, 2013. doi:10.1186/1471-2202-14-25.

Kewley-Port D, **Watson CS.** Formant-frequency discrimination for isolated English vowels. *J Acoust Soc Am* 95: 485–496, 1994. doi:10.1121/1.410024.

Lametti DR, **Nasir SM**, **Ostry DJ.** Sensory preference in speech production revealed by simultaneous alteration of auditory and somatosensory feedback. *J Neurosci* 32: 9351–9358, 2012. doi:10.1523/JNEUROSCI.0404-12.2012.

MacDonald EN, **Goldberg R**, **Munhall KG.** Compensations in response to real-time formant perturbations of different magnitudes. *J Acoust Soc Am* 127: 1059–1068, 2010. doi:10.1121/1.3278606.

Macmillan NA, **Creelman CD.** *Detection Theory: A User's Guide*. Mahwah, NJ: Lawrence Erlbaum Associates, 2004.

Munhall KG, **MacDonald EN**, **Byrne SK**, **Johnsrude I.** Talkers alter vowel production in response to real-time formant perturbation even when instructed not to compensate. *J Acoust Soc Am* 125: 384–390, 2009. doi:10.1121/1.3035829.

Nasir SM, **Ostry DJ.** Somatosensory precision in speech production. *Curr Biol* 16: 1918–1923, 2006. doi:10.1016/j.cub.2006.07.069.

Nasir SM, **Ostry DJ.** Speech motor learning in profoundly deaf adults. *Nat Neurosci* 11: 1217–1222, 2008. doi:10.1038/nn.2193.

Purcell DW, **Munhall KG.** Compensation following real-time manipulation of formants in isolated vowels. *J Acoust Soc Am* 119: 2288–2297, 2006. doi:10.1121/1.2173514.

Ranasinghe KG, **Kothare H**, **Kort N**, **Hinkley LB**, **Beagle AJ**, **Mizuiri D**, **Honma SM**, **Lee R**, **Miller BL**, **Gorno-Tempini ML**, **Vossel KA**, **Houde JF**, **Nagarajan SS.** Neural correlates of abnormal auditory feedback processing during speech production in Alzheimer's disease. *Sci Rep* 9: 5686, 2019. doi:10.1038/s41598-019-41794-x.

Rochet-Capellan A, **Ostry DJ.** Simultaneous acquisition of multiple auditory-motor transformations in speech. *J Neurosci* 31: 2657–2662, 2011. doi:10.1523/JNEUROSCI.6020-10.2011.

Romo R, **de Lafuente V.** Conversion of sensory signals into perceptual decisions. *Prog Neurobiol* 103: 41–75, 2013. doi:10.1016/j.pneurobio.2012.03.007.

Savariaux C, **Perrier P**, **Orliaguet JP.** Compensation strategies for the perturbation of the rounded vowel [u] using a lip tube: a study of the control space in speech production. *J Acoust Soc Am* 98: 2428–2442, 1995. doi:10.1121/1.413277.

Sidarta A, **van Vugt FT**, **Ostry DJ.** Somatosensory working memory in human reinforcement-based motor learning. *J Neurophysiol* 120: 3275–3286, 2018. doi:10.1152/jn.00442.2018.

Sperling G. The information available in brief visual presentations. *Psychol Monogr* 74: 1–29, 1960. doi:10.1037/h0093759.

Tourville JA, **Reilly KJ**, **Guenther FH.** Neural mechanisms underlying auditory feedback control of speech. *Neuroimage* 39: 1429–1443, 2008. doi:10.1016/j.neuroimage.2007.09.054.

Tremblay S, **Houle G**, **Ostry DJ.** Specificity of speech motor learning. *J Neurosci* 28: 2426–2434, 2008. doi:10.1523/JNEUROSCI.4196-07.2008.

Tremblay S, **Shiller DM**, **Ostry DJ.** Somatosensory basis of speech production. *Nature* 423: 866–869, 2003. doi:10.1038/nature01710.