UNIVERSITY OF BIRMINGHAM

The perils of learning to move while speaking

Lametti, Daniel R.; Quek, Marcus Y.M.; Prescott, Calum B.; Brittain, John Stuart; Watkins, Kate E.

Link to publication on Research at Birmingham portal

The perils of learning to move while speaking: One-sided interference between speech and visuomotor adaptation

Daniel R. Lametti, [1✉],[2]

Email danielrlametti@gmail.com

Marcus Y. M. Quek, [2]

Calum B. Prescott, [2]

John-Stuart Brittain, [3]

Kate E. Watkins, [2]

[1] Department of Psychology, Acadia University, Wolfville, Nova Scotia, Canada

[2] Department of Experimental Psychology, University of Oxford, Oxford, UK

[3] School of Psychology, University of Birmingham, Birmingham, UK

Abstract

Our understanding of the adaptive processes that shape sensorimotor behavior is largely derived from studying isolated movements. Studies of visuomotor adaptation, in which participants adapt cursor movements to rotations of the cursor's screen position, have led to prominent theories of motor control. In response to changes in visual feedback of movements, explicit (cognitive) and implicit (automatic) learning processes adapt movements to counter errors. However, movements rarely occur in isolation. The extent to which explicit and implicit processes drive sensorimotor adaptation when multiple movements occur simultaneously, as in the real world, remains unclear. Here we address this problem in the context of speech and hand movements. Participants spoke in-time with rapid, hand-driven cursor movements. Using real-time alterations of vowel sound feedback, and visual rotations of the cursor's screen position, we induced sensorimotor adaptation in one or both movements simultaneously. Across three experiments (n = 60, n = 48 and n = 76, respectively), we demonstrate that visuomotor adaptation is markedly impaired by simultaneous speech adaptation, and the impairment is specific to the explicit learning process in visuomotor adaptation. In contrast, visuomotor adaptation had no impact on speech adaptation. The results demonstrate that the explicit learning process in visuomotor adaptation is sensitive to movements in other motor domains. They suggest that some forms of speech adaptation may lack an explicit learning process altogether.

Keywords

Speech production
Adaptation
Visuomotor adaptation
Sensorimotor adaptation

Electronic supplementary material

The online version of this article (https://doi.org/10.3758/s13423-020-01725-8) contains supplementary material, which is available to authorized users.

Introduction

Motor behaviors rarely occur in isolation. We have conversations while sipping coffee, talk on the phone while walking down the street, and use gesticulations to facilitate speech. Even though we often make multiple movements simultaneously, research has largely focused on how motor behaviors are learned and maintained in isolation. For such an approach to yield broad theories of sensorimotor control, we must assume, firstly, that the control processes that govern one system (e.g., reaching) apply to other behaviors (e.g., speech), and, secondly, that these processes apply when multiple actions occur at the same time as they do in the real world. Here, we test these assumptions in speech and coincident hand movements.

Our understanding of sensorimotor control has been shaped by observing how participants adapt movements to perturbations of the sensory signals that guide movement. For example, sensorimotor adaptation has been extensively studied in the context of visual alterations of reaching movements. Classic work examined visuomotor adaptation related to prism-induced perturbations of vision (von Helmholtz, 1867). Contemporary studies of visuomotor adaptation have participants use a robotic arm, joystick, or stylus to move a cursor to targets on a computer screen (Cunningham, 1989; Krakauer et al., 2000; McDougle et al., 2015; Panouillères, Joundi, et al., 2015). After a series of baseline movements, rotations of the cursor's screen position are introduced so that movement towards a target at 90° appears to move along a 135° trajectory, movement towards a 135° target appears to move at 180° trajectory, and so on. Such visuomotor rotations create a mismatch between predicted and actual movement outcomes called a sensory prediction error. Over many trials, participants adapt their movements so that the cursor once again moves veridical to targets. This adaptation process, which occurs without conscious effort, involves updating internal representations that model the sensory consequences of reaching (Kawato, 1999; Wolpert et al., 2011).

Recent work suggests that visuomotor adaptation is also shaped by explicit (conscious) processes related to maintaining task goals (e.g., hitting a dartboard with a dart). When participants first encounter a large movement error, they employ aiming strategies that lead to rapid performance improvements (Bond & Taylor, 2015). If a cursor movement tracks a line at 45° to the left of an intended trajectory, participants aim 45° to the right on the next trial. These aiming strategies eventually diminish as the implicit process, working in parallel, drives the acquisition of more accurate motor plans (Mazzoni & Krakauer, 2006). Thus, reaching is maintained by two processes: an explicit process focused on maintaining task goals, and an implicit process focused on minimizing sensory prediction errors (McDougle et al., 2015, 2016; Taylor et al., 2014). These two processes double dissociate in behavior and the brain (Joseph M. Galea et al., 2011; Panouillères, Miall, et al., 2015; Schlerf et al., 2012; Taylor & Ivry, 2014; Tseng et al., 2007).

An experimental model akin to visuomotor adaptation has been used to study sensorimotor adaptation in speech. Participants produce speech into a microphone while the sound of their voice is played back to them in real-time through headphones (Houde & Jordan, 1998, 2002; Lametti et al., 2018). Using speech-processing software/hardware, the spectral properties that define vowel sounds (known as formants) are altered and played back to participants with an unnoticeable delay. Participants produce words such as "head" and "bed" into the microphone, but, because of the formant alteration, hear themselves producing different words (e.g., "had" and "bad"). Like visuomotor rotations, this speech perturbation creates a mismatch between the predicted and actual sensory consequences of movement. Over several hundred utterances, participants adapt their formant productions to reduce the perceived auditory sensory prediction error (Purcell & Munhall, 2006). This form of sensorimotor adaptation is thought to rely on the same implicit process observed in visuomotor adaptation (Houde & Nagarajan, 2011; Lametti et al., 2017; Tourville & Guenther, 2011). However, the extent to which speech adaptation involves a more explicit process remains unclear.

Few studies have examined the processes that govern sensorimotor adaptation when multiple movements with distinct goals are performed simultaneously. Such a test is important in the context of speech and limb movements, which frequently co-occur. We developed a dual-task paradigm to pair visuomotor adaptation of hand movements with adaptation to altered auditory feedback in speech. The task had participants use a handheld joystick to move a cursor into targets on a computer screen. In time with their hand movements, they produced words into a microphone and heard themselves in real-time over headphones. Here, we use this experimental model to induce sensorimotor learning in one or both movements simultaneously. In Experiment 1, we look for interactions between the processes that shape adaptation in each motor domain, and observe one-

sided interference. In Experiment 2 we replicate this result. In Experiment 3, we test whether the interference observed in Experiments 1 and 2 relates to the explicit component of visuomotor adaptation.

Methods

Participants

We recruited 184 right-handed, native English speakers, aged 18–35 years, from the University of Oxford community. Sixty participants (24 males) took part in Experiment 1, 48 (15 males) took part in Experiment 2, and 76 (30 males) took part in Experiment 3. Group sizes of 16–20 participants were chosen based on prior work in which significant between-group differences were observed with similar or smaller groups (Lametti et al., 2017; Panouillères, Joundi, et al., 2015). All participants were right-handed, had normal vision, hearing, and speech, and gave informed consent. The experiments were approved by the Central University Research Ethics Committee (R36869/RE003).

Apparatus

Participants sat at a table in front of a computer screen, 35 cm away at eye level. They grasped a joystick on the table with their right thumb and index finger. The joystick was 6.5 cm tall, had a center-out excursion of 17°, and was self-centering (APEM 9000 Series, RS Components). A shield hid participants' view of their joystick movements, which were sampled at 100 Hz. Participants wore a head-mounted microphone (Shure, WH20). Their speech was recorded at 44,100 Hz, mixed with 60 dB pink noise, and played back to them with an 11-ms delay through headphones (Sennheiser, HD 280 Pro). The formant structure of vowels was manipulated using acoustical effects processors (TC-Helicon, VoiceOne) and analog filters (Rockland, Wavetek) (Rochet-Capellan & Ostry, 2011). PsychoPy (www.psychopy.org) displayed stimuli on the screen and recorded speech. Electromyography (EMG) was recorded from a subsample of participants. Surface electrodes (ABRO, 22*30 mm) were placed on the right hand (first dorsal interosseous muscle) and the right side of the lips (orbicularis oris muscle). A ground electrode was placed on the wrist. EMG signals were sampled at 1,000 Hz (Cambridge Electronic Design, 1401).

Dual-task condition

The joystick controlled a red circle (diameter: 0.3 cm) that was in the center of the screen when the joystick was centered. At the start of each trial, a green target circle (diameter: 0.3 cm) appeared at one of eight equidistant positions, 4.6 cm from the screen's centre (see Fig. 1A). Each target position appeared once every eight trials (order randomized). Participants were instructed to move the red circle into the green circle while simultaneously producing the word "bed." They had 1,000 ms to make the joystick movement and say "bed." The inter-trial interval was 1,500 ms and during this time participants were instructed to let the joystick re-center. The cursor (the red circle) was visible throughout the experiment.

Fig. 1

Task and methods. (**A**) Participants used a joystick to move a cursor to eight equidistant targets arranged around a circle. The word "bed" was produced in strict time with their joystick movements. (**B**) Feedback was given at trial end to indicate if the peak velocity of the joystick movement (top panel) occurred at the same time as speech (bottom panel). "Move Sooner" was displayed if speech occurred before peak movement velocity; "Just Right" if speech occurred at the same time as peak movement velocity; and "Move Later" if speech occurred after peak movement velocity. (**C**) Electromyography was recorded from the hand and the lips in 14 participants to verify that the movements occurred simultaneously. (**D**) Speech adaptation (left panel) was induced by increasing the first formant frequency (F1) of perceived vowel sounds by 150 Hz (utterances 81–320, light blue circles). Visuomotor adaptation (right panel) was induced by rotating the cursor's screen position by 45° (movements 81–320, light-red circles)
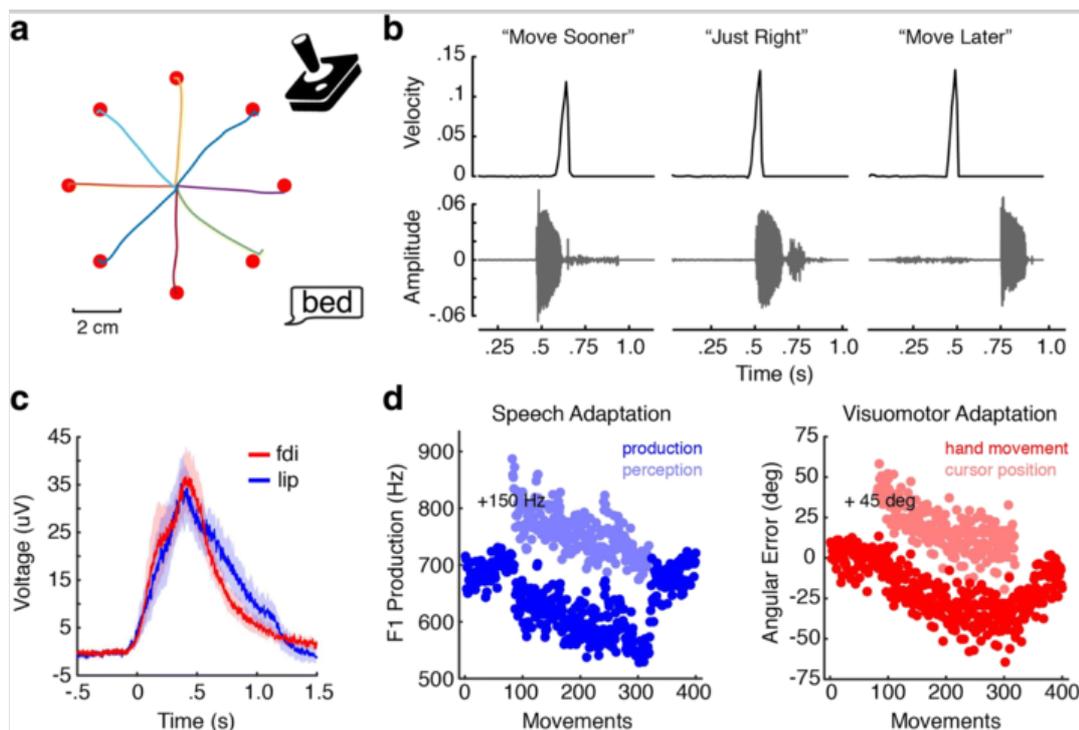
Feedback was displayed at trial end to indicate if the two movements occurred at the same time. "Just right!" was displayed if the peak velocity of the joystick movement occurred within 100 ms of "bed" being articulated; "Move sooner" was displayed if the peak velocity of the joystick movement occurred more than 100 ms after "bed"; "Move later" was displayed if the peak velocity of the joystick movement occurred more than 100 ms before "bed" (see Fig. 1B). Based on prior work, articulation was estimated to occur 50 ms before the center of the speech waveform (Lametti et al., 2012). "No response" was displayed if participants failed to move the joystick or speak. Although feedback referenced the joystick movements, participants were told that "Move sooner" was equivalent to "Speak later" and "Move later" to "Speak sooner." The average participant received "Just right!" feedback on 62% of trials. To verify that the feedback led participants to produce hand movements and speech simultaneously, muscle activity was recorded from the right hand and the lips in 14 participants. When data from all trials was included (i.e., "Move later," "Move sooner," "Just right!"), EMG activity associated with joystick movements and speech overlapped (see Fig. 1C). For this reason, data from all trials were included in the final analysis.

Single-task condition

A group of participants performed the joystick and speech production tasks on their own. For the joystick task, participants simply performed the same task described above without speaking. For the speech task, "bed" appeared on the screen in place of the target circle to prompt speech. The only feedback given was "no response" if participants failed to move the joystick or speak.

General procedure

Participants were given 40 practice trials. All experiments then had three phases: baseline, adaptation, and de-adaptation. Following 80 baseline movements, alterations to the cursor's screen position, the vowel sound in "bed," or both were applied to induce sensorimotor learning in one or both motor domains. To induce visuomotor adaptation, cursor movements were rotated by 45° in a counterclockwise direction relative to joystick movements. To induce speech adaptation, the first formant (F1) frequency of the vowel sound in productions of "bed" was increased by an average of 150 Hz from a baseline F1 frequency that averaged 696 Hz. Based on our past work (Lametti et al., 2014), the applied F1 alteration was large enough to place the vowel sound that participants heard into a new category such that participants produced "bed" but heard themselves producing a word that sounded more like "bad" (see Fig. 1D). These perturbations were applied for 240 trials (adaptation) and then removed for 80 trials (de-adaptation). The perturbations were either applied in full (Experiments 1 and 2) or they were built up in ten equal steps over the first 80 adaptation trials (Experiment 3). Participants were given a short break every 80 trials.

Data analysis

For the joystick task, the dependent measure was the angular error at peak movement velocity between the cursor's trajectory and a straight line from the start position to the target. Measures of angular error greater than 3 SD from the mean were excluded (< 1.5% of the data). Joystick adaptation was quantified as the difference in angular error between baseline movements and movements during adaptation. Participants adapted to the visuomotor alteration if the angular error of their joystick movements changed to oppose the 45° cursor rotation. For the speech task, linear predictive coding was used to calculate the mean F1 frequency based on a 50-ms segment at the center of each vowel. F1 values greater than 3 SD from the mean were excluded (< 1.5% of the data). Speech adaptation was quantified as the difference between baseline productions and productions during adaptation. Participants adapted to the speech perturbation if their F1 productions decreased to oppose the F1 increase that they heard.

We computed movement-onset times for both hand movements and speech production for all participants in the study by following a method described in Haith et al. (2015). For hand movements, position data were smoothed using a second-order Savitzky–Golay filter with a frame length of 11 samples. The data were differentiated and tangential velocity was calculated. Movement onset was identified as the time after trial start at which tangential velocity first exceeded 2.5 cm/s (less than 4% of peak velocity). Similarly, speech waveforms were high-pass filtered at 100 Hz and smoothed using a second-order Savitzky–Golay filter with a frame length of 11 samples. Speech onset was identified as the time after trial start at which the sound pressure level (SPL) of the waveform first exceeded .003 (less than 4% of peak SPL). In dual-task conditions, hand-movement onset averaged 520 ms and speech onset averaged 587 ms. When the tasks were performed on their own they averaged 469 ms and 525 ms, respectively. Movement onsets are displayed in Supplemental Table 1.

For hand movements, movement durations were calculated as the time during trials for which tangential hand velocity exceeded 2.5 cm/s. Movement durations averaged 211 ms in dual-task conditions, and 233 ms when the task was performed on its own. For speech production, durations were calculated as the time over which the speech waveform exceeded .003 SPL. Production durations averaged 309 ms in the dual-task conditions and 305 ms in the single-task condition.

Data and statistical analyses were performed in Matlab (Mathworks, Natick, MA, USA) and SPSS (IBM, Chicago, IL, USA). Changes in joystick angular error and changes in produced F1 frequency from baseline were calculated per subject and used as measures of adaptation. For each motor domain, repeated-measures analysis of variance and post hoc tests (two-tailed t-tests) were used to examine differences in adaptation and adaptation-related after-effects between groups. For multiple comparisons, the family-wise error rate was maintained using the Holm-Bonferroni method. Effect sizes (Cohen's d) were computed as the mean difference divided by the pooled standard deviation. Correlations were computed to explore the relationship between visuomotor and speech adaptation as well as speech and hand-movement onset times.

Results

Experiment 1 tested for interference between visuomotor adaptation and adaptation to altered auditory feedback. Sixty participants were divided into three dual-task conditions (Fig. 2A): (1) perturbed joystick movement and perturbed speech (n = 20), (2) perturbed joystick movement and normal speech (n = 20), and (3) normal joystick movement and perturbed speech (n = 20). In each group, the onset of hand movement correlated significantly with the onset of speech (r > .6, p < .001 in each case), suggesting that the two movements occurred at the same time.

Fig. 2

Experiment 1. (**A**) Group VS (red) experienced simultaneous visuomotor and speech adaptation. Group Vs (blue) experienced visuomotor adaptation while producing unaltred speech. Group vS (black) experienced speech adaptation while making unperturbed hand movements. (**B**) Left panel: Changes in F1 (in blocks of 16); trial numbers are on the x-axis. Increases in heard F1 drove reductions in produced F1 (grey area). Solid lines indicate the group mean and shading the standard error. Right panel: Average change in F1 over the adaptation phase of the experiment (movements 81–320). The points represent individual participants and the bars indicate the group mean with standard errors. Group Vs (blue) did not experience an F1 alteration. (**C**) Left panel: Changes in angular error (in blocks of 16); trial numbers are on the x-axis. A 45° rotation of the cursor drove changes in angular error (grey

area). Solid lines indicate the group mean and shading the standard error. Right panel: Average change in angular error over the adaptation phase of the experiment (movements 81–320). The points represent individual participants and the bars indicate the group mean with standard errors. The asterisk represents a significant difference between the indicated groups. Group vS (black) did not experience a cursor rotation
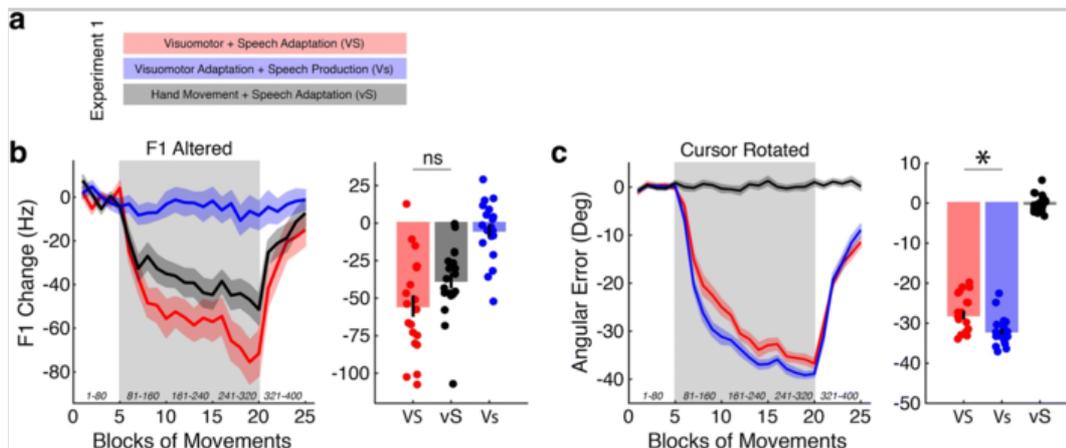

Figure 2B shows patterns of F1 change for these three groups. The right side of the figure shows average changes in F1 compared to baseline over the adaptation phase of the experiment (trials 81–320). Increases in perceived F1 drove compensatory changes in produced F1 compared to producing speech without altered feedback (red and black vs. blue data) ($F_{(2,57)} = 14.59$, $p < .001$). Compensatory changes in F1 production did not differ between the group who experienced speech and visuomotor adaptation simultaneously, and the group who experienced speech adaptation while simply making unperturbed hand movements (red vs. black data) ($t(38) = -1.89$, $p = .07$; Cohen's $d = .60$, CI = -.3 - 1.49). This lack of a difference in adaptation was reflected in the first eight after-effect trials (trials 321 to 328) that followed altered speech feedback ($t(38) = 1.55$, $p = .12$; Cohen's $d = .50$, CI = -.39 - 1.39). Thus, concurrent visuomotor adaptation did not alter speech adaptation.

Figure 2C shows changes in joystick angular error for the three groups in Experiment 1. The right side of the figure shows average changes in angular error over the adaptation phase of the experiment (trials 81–320). The cursor rotation drove compensatory changes in angular error compared to moving the joystick without a cursor rotation (red and blue vs. black data) ($F_{(2,57)} = 448.53$, $p < .001$). Compensation differed between the group who experienced simultaneous visuomotor and speech adaptation, and the group who experienced visuomotor adaptation while producing unperturbed speech (red vs. blue data) ($t(38) = 3.01$, $p = .005$; Cohen's $d = .95$, CI = .03 - 1.88). Concurrent speech adaptation reduced the amount of visuomotor adaptation. This difference in adaptation was also observed in the first eight after-effect trials (trials 321 to 328) that followed the restoration of veridical cursor feedback ($t(38) = 2.98$, $p = .005$; Cohen's $d = .94$, CI = .02 - 1.86). To test whether the reduction in visuomotor adaptation caused by speech adaptation reflected a trade-off with speech (i.e., as participants adapt more in one domain they adapt less in the other), we calculated the correlation between visuomotor and speech adaptation for the group who experienced both simultaneously. Visuomotor adaptation was not associated with speech adaptation ($r = -.22$, $p = .35$).

Experiment 1 suggests that altered speech feedback can interfere with concurrent visuomotor adaptation. Experiment 2 aimed to replicate this result, and to compare sensorimotor adaptation in the dual-task condition to sensorimotor adaptation in each motor domain alone. Forty-eight new participants were divided into two dual-task conditions and one condition in which the two tasks were experienced consecutively (Fig. 3A): (1) perturbed joystick movement and perturbed speech (n = 16), (2) perturbed joystick movement and normal speech (n = 16), and (3) perturbed joystick movement followed by perturbed speech, or vice versa (n = 16; task order counterbalanced). In the two dual-task conditions the onset of hand movement correlated significantly with the onset of speech across participants ($r > .6$, $p < .01$ in each case). This correlation was not observed when the tasks were performed consecutively ($r = .3$, $p = .25$).

Fig. 3

Experiment 2. (**A**) Group VS (red) experienced simultaneous visuomotor and speech adaptation. Group Vs (blue) experienced visuomotor adaptation while producing speech. Group S/V (black) experienced speech adaptation and

visuomotor adaptation consecutively. (**B**) Left panel: Changes in F1 (in blocks of 16): trials numbers are on the x-axis. Increases in perceived F1 drove reductions in produced F1 (grey area). Right panel: Average change in F1 over the adaptation phase of the experiment (movements 81–320). The points represent individual participants and the bars indicate the group mean with standard errors. Group Vs (blue) did not experience an F1 perturbation. (**C**) Left panel: Changes in angular error (in blocks of 16): trials numbers are on the x-axis. A 45° rotation of the cursor's screen position drove compensatory changes in angular error (grey area). Right panel: Average change in angular error over the adaptation phase of the experiment (movements 81–320). The points represent individual participants and the bars indicate the group mean with standard errors. The asterisk represents a significant difference between the indicated groups
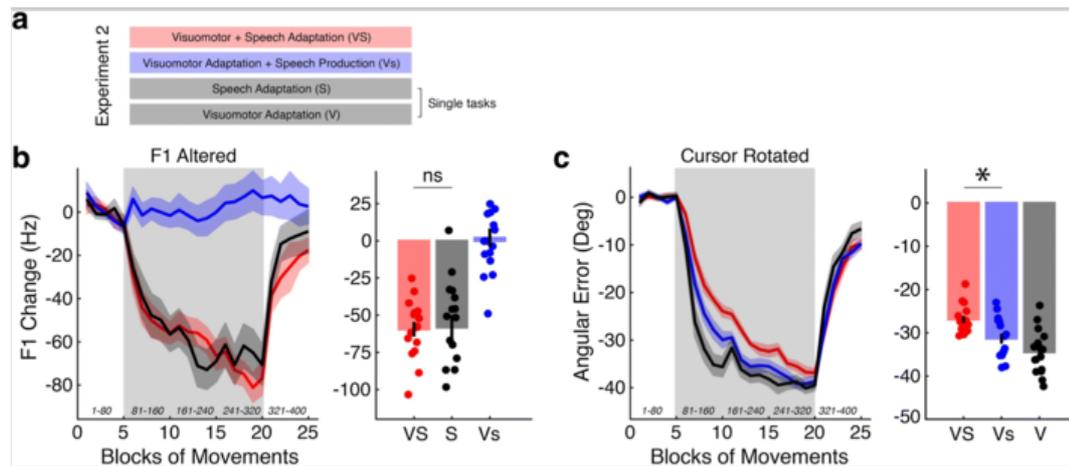

Figure 3B shows patterns of F1 change for the three groups in Experiment 2. The right side of the figure shows the average F1 change over the adaptation phase of the experiment (trials 81 to 320) compared to baseline. Increases in perceived F1 drove compensatory changes in produced F1 compared to simply producing speech without altered feedback (red and black vs. blue data) ($F(2,45) = 17.79$, $p < .001$). Critically, speech adaptation performed in isolation (i.e., without concurrent hand movements) was equivalent to speech adaptation with simultaneous visuomotor adaptation (black vs. red data) ($t(30) = .13$, $p = .9$; Cohen's $d = .05$, CI = -.83 - .92). This lack of a difference was mirrored in the first eight after-effect trials that followed altered speech feedback ($t(30) = .23$, $p = .8$; Cohen's $d = .08$, CI = -.80 - .96). Thus, speech adaptation was not altered by coincident hand movements, even when they involved visuomotor adaptation.

Figure 3C shows changes in joystick angular error associated with visuomotor adaptation for the three groups in Experiment 2. The right side of the figure shows average changes in angular error over the course of the experiment (trials 81–320). The presence of simultaneous speech production/adaptation drove significant between-group differences in the amount of visuomotor adaptation ($F(2,45) = 3.33$, $p < .05$). Replicating the main result in Experiment 1, altered auditory feedback reduced the amount of visuomotor adaptation compared to experiencing visuomotor adaptation while producing unperturbed speech (red vs. blue data) ($t(30) = 3.10$, $p = .004$; Cohen's $d = 1.10$, CI = .16 - 2.04). Unlike Experiment 1, a difference in adaptation-related after-effects in the eight trials that followed the restoration of veridical cursor feedback (trials 321–328) was not observed ($t(30) = .88$, $p = .38$; Cohen's $d = .31$, CI = -.57 - 1.2). Across the entire adaptation phase of the experiment, visuomotor adaptation while producing unaltered speech was not different to visuomotor adaptation on its own (blue vs. black data) ($t(30) = 1.79$, $p = .08$; Cohen's $d = .64$, CI = -.26 - 1.5). Finally, as in Experiment 1, the correlation between visuomotor adaptation and speech adaptation for the group who experienced both simultaneously was not significant ($r = .46$, $p = .07$). Reductions in visuomotor adaptation were due to the presence of altered speech feedback, as opposed to a tradeoff with speech adaptation.

Experiments 1 and 2 provide evidence that visuomotor adaptation is reduced by concurrent speech adaptation. One possibility is that concurrent speech adaptation interferes with the explicit component of visuomotor adaptation. Experiment 3 tested this hypothesis. All participants in Experiment 3 produced speech while making hand movements. To reduce the use of aiming strategies, visuomotor adaptation was driven by a gradual rotation of the cursor's screen position. Prior work suggests that eliminating the presence of large movement errors

reduces the use of aiming strategies as participants are less aware of a need to adapt (Butcher et al., 2017; Kagerer et al., 1997; Klassen et al., 2005).

Seventy-six participants were divided into four dual-task conditions in Experiment 3 (Fig. 4A): (1) perturbed joystick movement and perturbed speech introduced gradually (n = 20), (2) perturbed joystick movement and perturbed speech in which only the joystick perturbation was introduced gradually (n = 16), (3) perturbed joystick movement and normal speech in which the joystick perturbation was introduced gradually (n = 20), and (4) normal joystick movements and perturbed speech in which the speech perturbation was introduced gradually (n = 20). The onset of hand movement correlated significantly with the onset of speech across participants (r > .57, p < .03 in each case).

Fig. 4

Experiment 3: (**A**) Group VS (red) experienced gradual visuomotor and speech adaptation. Group Vs (blue) experienced gradual visuomotor adaptation while producing unperturbed speech. Group vgS (black) experienced gradual speech adaptation while making unperturbed hand movements. Group gVS (grey) experienced abrupt speech adaptation and gradual visuomotor adaptation. (**B**) Left panel: Changes in F1 (in blocks of 16): trials numbers are on the x-axis. Increases in heard F1 drove reductions in produced F1 (grey area). Right panel: Average change in F1 when the F1 alteration was applied in full (movements 161 to 320). The points represent individual participants and the bars indicate the group mean with standard errors. Group gVs (blue) did not experience an F1 alteration. (**C**) Left panel: Changes in angular error (in blocks of 16) over the course of the experiment. A 45° rotation of the cursor's screen position was built up between trials 81 and 160 to drive gradual changes in angular error (grey area). Right panel: Average change in angular error when the cursor rotation was applied in full (movements 161–320). The points represent individual participants and the bars indicate the group mean with standard errors. The error bars represent standard errors. Group vgS (black) did not experience a cursor rotation
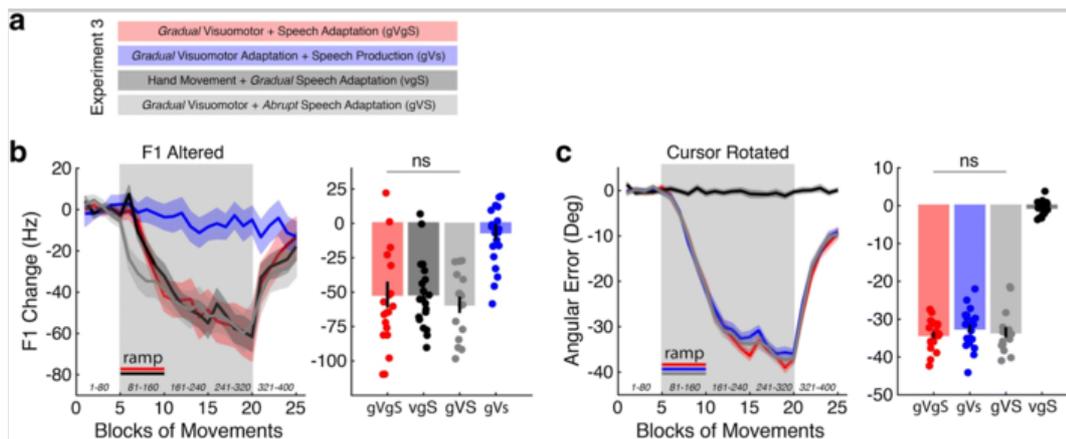

Figure 4B shows patterns of F1 change for these four groups. The right side of the figure shows the average F1 change over the last 160 trials of the adaptation phase of the experiment in which the speech alteration was fully introduced for all groups (trials 161–320). Increases in perceived F1 drove compensatory changes in produced F1 (red, grey, and black data) compared to simply producing speech without altered feedback (blue data) $(F(3,72) = 7.44, p < .001)$. There were no differences in the amount of speech adaptation between the groups who made coincident hand movements or simultaneously learned a visuomotor adaptation (red vs. grey vs. black data) $(p > .3$ in each case; Cohen's d < .31 in each case).

Figure 4C shows changes in angular error associated with hand movements for the four groups in Experiment 4. The right side of the figure shows the average change in angular error over the last 160 trials of the adaptation phase of the experiment in which the visuomotor alteration was fully introduced (trials 161–320). The gradual introduction of the 45° cursor rotation drove compensatory changes in angular error (red, grey, and blue data) compared to simply moving the joystick without a cursor rotation (black data) $(F(3,72) = 302.31, p < .001)$. When the cursor rotation was introduced gradually the presence of coincident speech adaptation did not alter the amount of visuomotor adaptation, regardless of whether speech adaptation was introduced gradually (red vs. blue data) $(t(38) = 1.2, p = .23$; Cohen's d = .38, CI = -0.50 - 1.27) or abruptly (grey vs. blue data) $(t(34) = .58,$

p = .57; Cohen's d = .19, CI = -0.69 - 1.07). This lack of a difference was mirrored in the first eight after-effect trials that followed altered visual feedback (p > .1 in each case; Cohen's d < .6 in each case). The impact of altered speech feedback on visuomotor adaptation in Experiments 1 and 2 depended on the presence of large movement errors at the start of adaptation. The results are consistent with the hypothesis that sensorimotor adaptation in speech interferes with the explicit component of visuomotor adaptation.

Discussion

Sensorimotor adaptation has largely been studied in the context of isolated movements. However, movements with different goals frequently co-occur. It remained unclear whether the processes that support sensorimotor adaptation in isolated movements shape adaptation during multi-movement behaviors. We tested this idea in the context of speech and hand movements. Participants produced speech while simultaneously making movements of a joystick to control a cursor on a computer screen. Real-time alterations in vowel sounds, and visuomotor rotations of the cursor's screen position, drove sensorimotor adaptation in one or both motor domains. In an initial experiment, we found that altered speech feedback reduced visuomotor adaptation. In a second experiment, we replicated this observation. In a final experiment, this interference disappeared when the visuomotor rotation was gradually applied to reduce explicit adaptation strategies. The results suggest that the explicit component of visuomotor adaptation is susceptible to interference from coincident movements that contain an error signal.

In both Experiment 1 and Experiment 2, we failed to find a correlation between speech adaptation and visuomotor adaptation. That is, interference between altered speech feedback and visuomotor adaptation did not reflect a trade-off between adaptation in each domain. Visuomotor adaptation was reduced by the mere presence of altered feedback. One explanation for this finding is that altered speech feedback taxed working memory, disrupting the explicit component of visuomotor adaptation (Christou et al., 2016). We expand on this idea below.

Like many studies of visuomotor adaptation, here participants rapidly moved a cursor to equidistant targets arranged around a circle (Bond & Taylor, 2015; Cunningham, 1989; Joseph M. Galea et al., 2011; Krakauer et al., 2000; Mazzoni & Krakauer, 2006; Panouillères, Joundi, et al., 2015). A visuomotor rotation altered visual feedback of cursor movements in a counterclockwise direction that varied with target direction. An upward movement on the screen to 0° resulted in a leftward alteration of the cursor's screen position, whereas a downward movement at 180° resulted in a rightward alteration (and so on). To strategically counter this visuomotor rotation with aiming movements, participants must hold in working memory the direction in which the cursor position was altered for each of the eight movement directions. A growing body of work suggests that visuomotor adaptation draws heavily on working memory (McDougle & Taylor, 2019). During the early stages of visuomotor adaptation, patterns of brain activity reflect those observed during spatial working-memory tasks (e.g., mental rotations of objects) (Anguera et al., 2010). Age-related declines in working memory have been linked to deficits in visuomotor adaptation (Anguera et al., 2011), and positive associations between working-memory capacity and the use of explicit aiming strategies in visuomotor adaptation have also been observed (Christou et al., 2016). Resource models of working memory posit that as task demands increase working-memory capacity declines (Ma et al., 2014). An additional strain on working memory from altered speech feedback might explain the reduction in visuomotor adaptation observed here (J. M. Galea et al., 2010). More fundamentally, the results suggest that strategy-based learning may play a minor role in complex motor tasks involving multiple effectors and error signals.

One limitation of this study is that we do not directly test if speech adaptation draws on explicit learning mechanisms. Nevertheless, the results suggest that adaptation to altered vowel sounds may have a reduced strategy-based component compared to visuomotor adaptation. Here, in all cases, adaptation to a formant alteration was unaffected by a second sensorimotor task. At least one other study supports the idea that sensorimotor adaptation in speech may have a reduced explicit learning component. Munhall et al. (2009) instructed two groups of participants experiencing formant manipulations to (1) ignore the sounds they heard from the headphones or (2) resist altering their speech by instead focusing on kinesthetic feedback. Despite these instructions, at every stage of sensorimotor learning their formant productions tracked adaptive changes in vowel production produced by naive controls. Likewise, implicit adaptation cannot be suppressed in visuomotor adaptation, even when participants are made aware of the alteration. However, in visuomotor adaptation,

additional explicit adaptation strategies seem to be more frequently employed, especially early in learning (Mazzoni & Krakauer, 2006).

It remains unknown whether strategic changes in speech production can be used to counter different types of feedback manipulations such as pitch changes. In these cases, participants may better understand how to change their speech to reduce perceived errors. Although participants are typically aware of large formant alterations, they may lack knowledge of how speech movements relate to vowel sounds. Training participants on the relationship between tongue position and vowel production prior to altered formant feedback could lead to the use of explicit adaptation strategies. More generally, the conditions under which participants use explicit learning mechanisms in speech adaptation should be explored in future work.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Acknowledgements

Open Practices Statement The reported experiments were not pre-registered. The raw data have not been made available on a third-party archive because speech recordings may reveal the identity of participants. The derived data from speech recordings, joystick movements, and all analysis code are available from the first author upon reasonable request.

Electronic supplementary material

ESM 1

(DOCX 19 kb)

References

Anguera, J. A., Reuter-Lorenz, P. A., Willingham, D. T., & Seidler, R. D. (2010). Contributions of spatial working memory to visuomotor learning. *Journal of Cognitive Neuroscience, 22*(9), 1917–1930. https://doi.org/10.1162/jocn.2009.21351

Anguera, J. A., Reuter-Lorenz, P. A., Willingham, D. T., & Seidler, R. D. (2011). Failure to engage spatial working memory contributes to age-related declines in visuomotor learning. *Journal of Cognitive Neuroscience*, *23*(1), 11–25.

Bond, K. M., & Taylor, J. A. (2015). Flexible explicit but rigid implicit learning in a visuomotor adaptation task. *Journal of Neurophysiology*, *113*(10), 3836–3849.

Butcher, P. A., Ivry, R. B., Kuo, S.-H., Rydz, D., Krakauer, J. W., & Taylor, J. A. (2017). The cerebellum does more than sensory prediction error-based learning in sensorimotor adaptation tasks. *Journal of Neurophysiology, 118*(3), 1622–1636. https://doi.org/10.1152/jn.00451.2017

Christou, A. I., Chris Miall, R., McNab, F., & Galea, J. M. (2016). Individual differences in explicit and implicit visuomotor learning and working memory capacity. *Scientific Reports, 6*(1). https://doi.org/10.1038/srep36633

Cunningham, H. A. (1989). Aiming error under transformed spatial mappings suggests a structure for visual-motor maps. *Journal of Experimental Psychology. Human Perception and Performance*, *15*(3), 493–506.

Galea, J. M., Sami, S. A., Albert, N. B., & Miall, R. C. (2010). Secondary tasks impair adaptation to step- and gradual-visual displacements. *Experimental Brain Research, 202*(2), 473–484. https://doi.org/10.1007/s00221-010-2158-x

Galea, J. M., Vazquez, A., Pasricha, N., de Xivry, J.-J. O., & Celnik, P. (2011). Dissociating the roles of the cerebellum and motor cortex during adaptive learning: the motor cortex retains what the cerebellum learns. *Cerebral Cortex*, *21*(8), 1761–1770.

Haith, A. M., Huberdeau, D. M., & Krakauer, J. W. (2015). The influence of movement preparation time on the expression of visuomotor learning and savings. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *35*(13), 5109–5117.

Houde, J. F., & Nagarajan, S. S. (2011). Speech production as state feedback control. *Frontiers in Human Neuroscience*, *5*, 82.

Houde, & Jordan. (1998). Sensorimotor adaptation in speech production. *Science*, *279*(5354), 1213–1216.

Houde, & Jordan. (2002). Sensorimotor adaptation of speech I: Compensation and adaptation. *Journal of Speech, Language, and Hearing Research: JSLHR*, *45*(2), 295–310.

Kagerer, F. A., Contreras-Vidal, J. L., & Stelmach, G. E. (1997). Adaptation to gradual as compared with sudden visuo-motor distortions. *Experimental Brain Research. Experimentelle Hirnforschung. Experimentation Cerebrale*, *115*(3), 557–561.

Kawato, M. (1999). Internal models for motor control and trajectory planning. *Current Opinion in Neurobiology*, *9*(6), 718–727.

Klassen, J., Tong, C., & Flanagan, J. R. (2005). Learning and recall of incremental kinematic and dynamic sensorimotor transformations. *Experimental Brain Research. Experimentelle Hirnforschung. Experimentation Cerebrale*, *164*(2), 250–259.

Krakauer, J. W., Pine, Z. M., Ghilardi, M. F., & Ghez, C. (2000). Learning of visuomotor transformations for vectorial planning of reaching trajectories. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 20*(23), 8916–8924.

Lametti, D. R., Krol, S. A., Shiller, D. M., & Ostry, D. J. (2014). Brief periods of auditory perceptual training can determine the sensory targets of speech motor learning. *Psychological Science, 25*(7), 1325–1336.

Lametti, D. R., Nasir, S. M., & Ostry, D. J. (2012). Sensory preference in speech production revealed by simultaneous alteration of auditory and somatosensory feedback. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *32*(27), 9351–9358.

Lametti, D. R., Smith, H. J., Freidin, P., & Watkins, K. E. (2017). Cortico-cerebellar Networks Drive Sensorimotor Learning in Speech. *Journal of Cognitive Neuroscience*, 1–12.

Lametti, D. R., Smith, H. J., Watkins, K. E., & Shiller, D. M. (2018). Robust Sensorimotor Learning During Variable Sentence Level Speech. *Current Biology: CB*.

Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of working memory. *Nature Neuroscience*, *17*(3), 347–356.

Mazzoni, P., & Krakauer, J. W. (2006). An implicit plan overrides an explicit strategy during visuomotor adaptation. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *26*(14), 3642–3645.

McDougle, S. D., Bond, K. M., & Taylor, J. A. (2015). Explicit and Implicit Processes Constitute the Fast and Slow Processes of Sensorimotor Learning. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 35*(26), 9568–9579.

McDougle, S. D., Ivry, R. B., & Taylor, J. A. (2016). Taking Aim at the Cognitive Side of Learning in Sensorimotor Adaptation Tasks. *Trends in Cognitive Sciences, 20*(7), 535–544.

McDougle, S. D., & Taylor, J. A. (2019). Dissociable cognitive strategies for sensorimotor learning. *Nature Communications, 10*(1), 40.

Munhall, K. G., MacDonald, E. N., Byrne, S. K., & Johnsrude, I. (2009). Talkers alter vowel production in response to real-time formant perturbation even when instructed not to compensate. *The Journal of the Acoustical Society of America, 125*(1), 384–390.

Panouillères, M. T. N., Joundi, R. A., Brittain, J.-S., & Jenkinson, N. (2015). Reversing motor adaptation deficits in the ageing brain using non-invasive stimulation. *The Journal of Physiology*, *593*(16), 3645–3655.

Panouillères, M. T. N., Miall, R. C., & Jenkinson, N. (2015). The role of the posterior cerebellum in saccadic adaptation: a transcranial direct current stimulation study. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *35*(14), 5471–5479.

Purcell, D. W., & Munhall, K. G. (2006). Adaptive control of vowel formant frequency: evidence from real-time formant manipulation. *The Journal of the Acoustical Society of America, 120*(2), 966–977.

Rochet-Capellan, A., & Ostry, D. J. (2011). Simultaneous acquisition of multiple auditory-motor transformations in speech. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *31*(7), 2657–2662.

Schlerf, J., Ivry, R. B., & Diedrichsen, J. (2012). Encoding of sensory prediction errors in the human cerebellum. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *32*(14), 4913–4922.

Taylor, J. A., & Ivry, R. B. (2014). Cerebellar and prefrontal cortex contributions to adaptation, strategies, and reinforcement learning. *Progress in Brain Research, 210*, 217–253.

Taylor, J. A., Krakauer, J. W., & Ivry, R. B. (2014). Explicit and implicit contributions to learning in a sensorimotor adaptation task. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *34*(8), 3023–3032.

Tourville, J. A., & Guenther, F. H. (2011). The DIVA model: A neural theory of speech acquisition and production. *Language and Cognitive Processes*, *26*(7), 952–981.

Tseng, Y.-W., Diedrichsen, J., Krakauer, J. W., Shadmehr, R., & Bastian, A. J. (2007). Sensory prediction errors drive cerebellum-dependent adaptation of reaching. *Journal of Neurophysiology*, *98*(1), 54–62.

von Helmholtz, H. (1867). *Handbuch der physiologischen Optik*. Voss.

Wolpert, D. M., Diedrichsen, J., & Flanagan, J. R. (2011). Principles of sensorimotor learning. *Nature Reviews. Neuroscience*, *12*(12), 739–751.