

1 Variability is actively regulated in speech

2
3 Ding-lan Tang¹, Benjamin Parrell^{1‡*}, Caroline A. Niziolek^{1‡*}

4 ¹Department of Communication Sciences & Disorders, University of Wisconsin–Madison,
5 Madison, WI, USA

6 ‡ these authors contributed equally to this work

7 * Correspondence: bparrell@wisc.edu, cniziolek@wisc.edu

9 Abstract

10 Although movement variability is often attributed to unwanted noise in the motor system, recent
11 work has demonstrated that variability may be actively controlled. To date, research on
12 regulation of motor variability has relied on relatively simple, laboratory-specific reaching tasks.
13 It is not clear how these results translate to complex, well-practiced and real-world tasks. Here,
14 we test how variability is regulated during speech production, a complex, highly over-practiced
15 and natural motor behavior that relies on auditory and somatosensory feedback. Specifically, in a
16 series of four experiments, we assessed the effects of auditory feedback manipulations that
17 modulate perceived speech variability, shifting every production either towards (*inward-pushing*)
18 or away from (*outward-pushing*) the center of the distribution for each vowel. Participants
19 exposed to the *inward-pushing* perturbation (Experiment 1) increased produced variability while
20 the perturbation was applied as well as after it was removed. Unexpectedly, the *outward-pushing*
21 perturbation (Experiment 2) also increased produced variability during exposure, but variability
22 returned to near baseline levels when the perturbation was removed. Outward-pushing
23 perturbations failed to reduce participants' produced variability both with larger perturbation
24 magnitude (Experiment 3) or after their variability had increased above baseline levels as a result
25 of the inward-pushing perturbation (Experiment 4). Simulations of the applied perturbations
26 using a state space model of motor behavior suggest that the increases in produced variability in
27 response to the two types of perturbations may arise through distinct mechanisms: an increase in
28 controlled variability in response to the inward-pushing perturbation, and an increase in
29 sensitivity to auditory errors in response to the outward-pushing perturbation. Together, these
30 results suggest that motor variability is actively regulated even in complex and well-practiced
31 behaviors, such as speech.

32
33
34
35
36
37
38

Keywords: motor variability, speech motor control, sensorimotor adaptation, auditory
perturbation, error sensitivity.

39 **Introduction**

40 No matter how hard we practice, it is virtually impossible to generate exactly the same
41 movement twice. Such variation in performance across repetitions of the same movements, or
42 *motor variability*, is widely believed to be an inevitable consequence of noise in the nervous
43 system, arising from stochastic events presented across all scales of brain activity, from the
44 single-cell level to complex network dynamics (Churchland et al., 2006; Faisal et al., 2008;
45 Renart & Machens, 2014). Indeed, many current theories of motor behavior rely on this
46 assumption and posit that the motor system aims to minimize the detrimental effects of ‘motor
47 noise’ on motor task performance (Harris & Wolpert, 1998; Scholz & Schöner, 1999; Todorov,
48 2004).

49
50 However, recent work in reaching has demonstrated that variability is not always treated as
51 unwanted “noise” to be reduced by the motor system but may be more actively controlled.
52 Repeated exposure to position- or velocity-dependent force fields during reaching has been
53 shown to selectively increase task-relevant variability, potentially to facilitate more efficient
54 future learning (Wu et al., 2014). Conversely, task-relevant variability can also be reduced when
55 needed in some behaviors: participants exposed to a visual perturbation that magnified the
56 horizontal displacement of the hand away from the midline during point-to-point reaching
57 movements reduced their variability in this dimension (Wong et al., 2009). These results suggest
58 that variability is not simply noise but can also be an important part of the signal itself that
59 controls the motor movement (Stein et al., 2005).

60
61 Although variability poses a fundamental problem for motor control, studies on regulation of
62 motor variability are comparatively sparse and, to date, have relied principally on relatively
63 simple, laboratory-specific planar arm reaching tasks. While such tasks play an important role in
64 probing motor control systems, they tackle a relatively restricted range of motor tasks that do not
65 fully capture the complex demands of real-world behavior and, thus, may overlook some aspects
66 of motor control in real-world tasks which are naturally less constrained. Speech production is
67 one such task: as opposed to laboratory-specific arm reaching tasks, which involve an interaction
68 with an uncommon external device (e.g. a joystick) and typically constrain movements to two
69 joints, speaking is a highly over-practiced behavior that involves the coordination of roughly 100
70 muscles to precisely control the respiratory and phonatory systems as well as the movements of
71 lips, jaw, velum and tongue (a muscular hydrostat with highly complex control).

72
73 Here, we aimed to test how variability is regulated in speech production, a complex, well-
74 practiced task controlled via non-visual sensory feedback. Previous work has shown that
75 speakers are sensitive to real-time alterations to their auditory feedback (Houde & Jordan, 1998;
76 Purcell & Munhall, 2006). For example, speakers learn to alter their speech to oppose auditory
77 perturbations that shift the frequencies of vowel formants (resonances of the vocal tract which
78 serve to distinguish between different vowels). However, these studies have typically examined

79 formant shifts that are applied in a consistent direction, regardless of the produced vowel
80 formants, with the consequence that the mean formant values participants hear are altered, while
81 their formant variability remains unchanged. Here, we implement a novel auditory perturbation
82 that shifts vowel formants in a non-uniform manner, such that the mean formant values remain
83 unchanged while trial-to-trial variability is either increased or decreased (Fig. 1).

84
85 By altering participants' perceived trial-to-trial variability without affecting their overall mean
86 behavior, we can test whether variability in speech production is actively monitored and
87 regulated. In a series of four experiments, we assess the effects of manipulations that both
88 increase and decrease the perceived variability of participant's speech behavior. We predicted
89 that a perturbation that reduces perceived variability would lead to increases in produced
90 variability, as participants would be free to be less precise in their production without negatively
91 affecting their perceived accuracy. Conversely, we expected that a perturbation that increases
92 perceived variability would have the opposite effect (i.e. lead to decreases in produced
93 variability). Surprisingly, we found that both types of perturbation caused participants to increase
94 their produced variability, while only the perturbation which increases perceived variability
95 affected a behavioral measure of error sensitivity. Simulations of the applied perturbations using
96 a well-established state space model of sensorimotor learning suggest that the increases in
97 produced variability in response to the two types of perturbations may arise through distinct
98 mechanisms: an increase in controlled variability in response to the perturbation that reduces
99 perceived variability, and an increase in sensitivity to auditory errors in response to the
100 perturbation that increases perceived variability. All together, these results suggest that motor
101 variability is actively monitored and regulated even in complex and well-practiced behaviors,
102 such as speech.

103

104 **Methods**

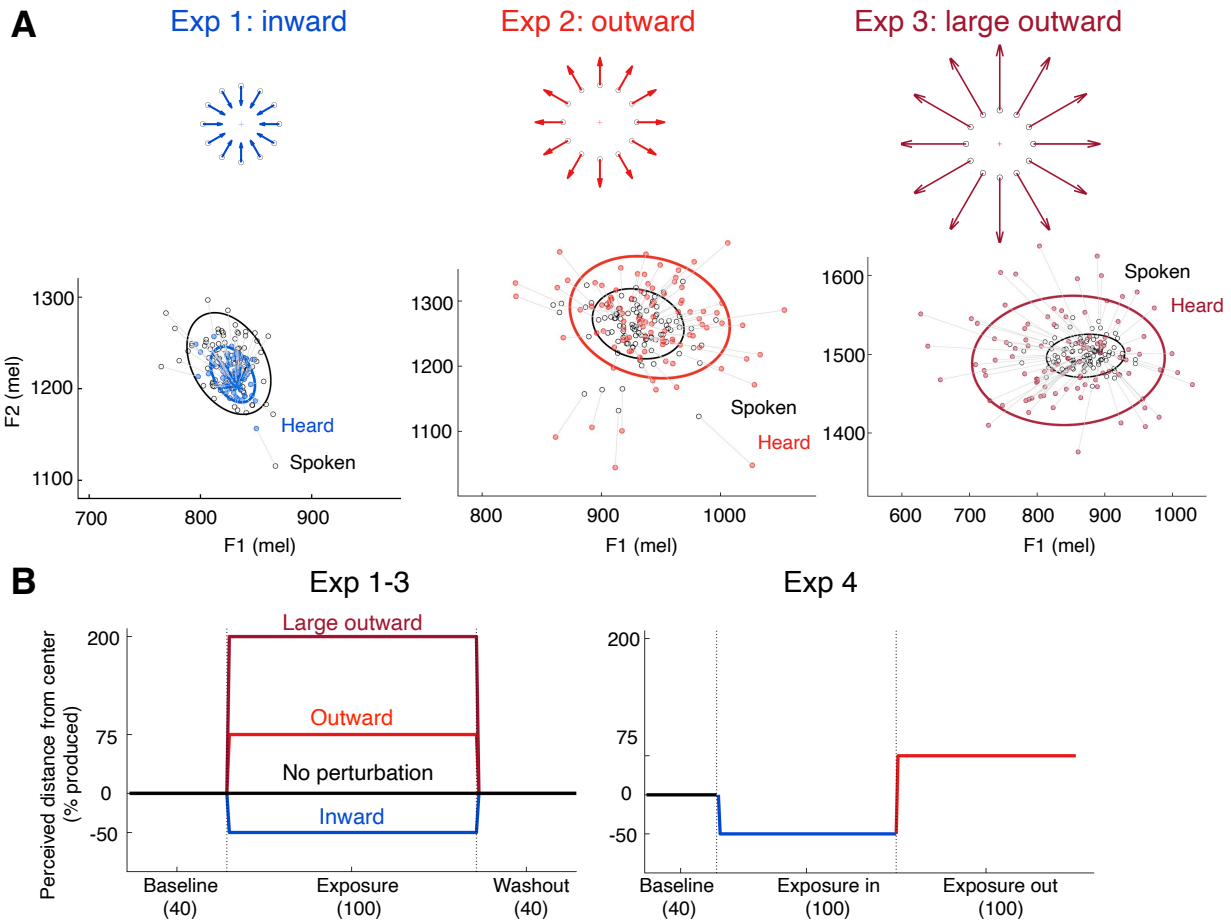
105 *Participants*

106 87 native speakers of American English between the ages of 18 and 66 years, with no reported
107 history of hearing loss or neurological disorders, took part in the study (Experiment 1: N = 24,
108 28.8 ± 12.3 years, 18/6 females/males; Experiment 2: N = 22, 25.9 ± 9.7 years, 14/8
109 females/males; Experiment 3: N = 21, 33.3 ± 13.6 years, 12/9 females/males; Experiment 4: N =
110 20, 28 ± 12.9 years, 15/5 females/males). Participants provided informed consent and were
111 compensated either monetarily or with course credit for their participation. The Institutional
112 Review Board of the University of Wisconsin–Madison approved the experimental protocol.

113

114

115



116

117

118 **Figure 1. Experiment design.** (A) Schematic (top) and examples from representative
 119 participants (bottom) of perturbations applied to speech vowel formants: inward-pushing (Exp 1,
 120 left), outward-pushing (Exp 2, middle), and large outward-pushing (Exp 3, right). Note different
 121 axis scales in the example data across experiments. Black circles represent the formant values
 122 participants produced and colored circles indicate what was played back to them over
 123 headphones. The ellipses represent a 95% confidence interval around the data points of the same
 124 color. (B) Experimental procedure and magnitudes of the perturbation applied in Experiments 1-
 125 3 (left panel) and Experiment 4 (right panel).

126

127 Apparatus

128 Participants were exposed to a real-time perturbation of the first and second vowel formants
 129 (F1/F2) designed to affect the perceived variability of their speech production. A modified
 130 version of Audapter (Cai et al., 2008; Tourville et al., 2013) was used to record participants'
 131 speech, alter the speech signal when necessary, and play the (potentially altered) signal back to
 132 participants. The experiment was conducted in a quiet room with participants seated in front of a
 133 computer screen. In each trial, one of the three stimulus words (“bead”, “bad”, and “bod” in

134 Experiment 1, 2, 4 or “bead”, “bad”, and “bed” in Experiment 3) was pseudorandomly selected
135 and displayed on the screen, and participants read it aloud. Speech was recorded at 16 kHz via
136 either a head-mounted microphone (AKG C520, Exp 1-2) or a desktop microphone (Sennheiser
137 MKE 600, Exp 3-4). The output of Audapter was played back to participants via closed-back
138 circumaural headphones (Beyerdynamic DT 770) with an unnoticeable delay of ~18 ms, as
139 measured on our system following Kim & Max (2020). All trials were processed through
140 Audapter in the same manner, regardless of whether a perturbation was applied. Speech was
141 played back at a volume of approximately 80 dB SPL and mixed with speech-shaped noise at
142 approximately 60 dB SPL, which served to mask potential perception of the participants’ own
143 unaltered speech through air and bone conduction. The volume of speech playback varied
144 dynamically with the amplitude of participants’ produced speech.
145

146 *Experiment-specific auditory perturbation*

147 We designed a modified version of Audapter (Parrell & Niziolek, 2021) that is able to affect the
148 perceived variability of speech production by specifying formant perturbations as a function of
149 the current values of F1 and F2 (Figure 1A). A participant-specific perturbation field was
150 calculated for each vowel such that every production during the exposure phase was shifted
151 toward (inward-pushing perturbation) or away from (outward-pushing perturbation) the mean
152 F1/F2 values of that participant’s distribution for that vowel (the vowel “targets”). The
153 magnitude of the perturbation was defined as a percentage of the distance between the currently
154 produced vowel formants and the vowel targets. By scaling the error between the vowel formants
155 and their targets, the inward- and outward-pushing perturbations respectively reduce and
156 magnify the perceived variability of speech production.
157

158 In Experiment 1 (**inward-pushing perturbation**), participants received a perturbation that
159 shifted every production *towards* the vowel target. The perturbation was 50% of the distance, in
160 F1/F2 space, between the current formant values and the vowel targets (see Figure S1).
161

162 In Experiment 2 (**outward-pushing perturbation**), participants received the opposite
163 perturbation, a shift of every production *away* from these targets, with a slightly larger
164 perturbation magnitude (i.e., 75% of the distance to the vowel target, Figure S1).
165

166 In Experiment 3 (**large outward-pushing perturbation**), we aimed to test the possibility that
167 the failure to find the hypothesized reduction in variability in response to the outward-pushing
168 perturbation applied in Experiment 2 was due to inadequate perturbation magnitude. We
169 suspected that a larger perturbation magnitude might be needed to drive participants to produce
170 compensatory reductions in variability. The perturbation magnitude was therefore increased to
171 200% of the distance to the vowel target in Experiment 3 (Figure S1) and one stimulus word was

172 changed from a corner vowel word (“bad”) to a non-corner vowel word (“bed”) to test the effect
173 of the perturbation on non-corner vowels.

174

175 In Experiment 4 (**inward-outward pushing perturbation**), we further examined whether limits
176 on articulatory precision prevented participants from reducing their produced variability in
177 response to outward-pushing perturbations in Experiment 2 and 3: that is, if individuals already
178 produce vowels at the lower limit of variability, they may not be able to produce further
179 variability decreases. In Experiment 4, participants experienced two exposure phases: first, to an
180 inward-pushing perturbation, which served to increase participants’ produced variability above
181 baseline levels, and then to an outward-pushing perturbation. The perturbation magnitudes were
182 50% of the distance to the vowel target for both inward-pushing and outward-pushing phases.

183

184 In order to control for potential changes in variability over the course of ~500 trials of single
185 word production, we additionally analyzed an existing dataset (Parrell & Niziolek, 2021) with a
186 similar experimental structure, though with no auditory perturbation applied (auditory feedback
187 was processed through Audapter in the same manner as the baseline phases in Experiments 1-4).
188 This control experiment included 460 trials of single word productions (115 repetitions of each
189 stimulus word in Experiments 1, 2 and 4 — “bead”, “bad”, and “bod”—as well as the word
190 “booed”, which did not occur in Experiments 1-4 and was not analyzed). To match the
191 experimental design of the current study, the 460 trials were divided into four phases: baseline
192 (30 trials per stimulus), early exposure (30 trials per stimulus), late exposure (30 trials per
193 stimulus) and washout (25 trials per stimulus). The experimental setup for the control dataset,
194 including recording, processing, and headphone presentation of speech, was identical to the
195 Experiments 1-4.

196

197 ***Procedure***

198 In all experiments, stimulus words were presented on the computer screen for 1.5 s, one at a
199 time. The interstimulus interval was randomly jittered between 0.25-1 s. Participants were
200 instructed to read each word out loud as it appeared.

201

202 Each experiment had three phases (Figure 1B). Experiments 1-3 were divided into baseline,
203 exposure and washout phases. In the baseline phase (40 trials per stimulus), participants received
204 unaltered auditory feedback and we measured participants’ mean F1/F2 values for each vowel.
205 These values were subsequently used to calculate the participant-specific perturbation field (see
206 above). The exposure phase followed the baseline phase. In the exposure phase, participants
207 produced each stimulus 100 times while receiving either the inward-pushing, outward-pushing,
208 or large outward-pushing perturbation (see experiment-specific auditory perturbation above).
209 Experiments 1-3 ended with a washout phase where participants produced each word 40 times
210 with unaltered auditory feedback. In Experiment 4, following the baseline phase, participants

211 experienced two sequential 300-trial (each stimulus 100 times) exposure phases, first with an
212 inward-pushing perturbation and then with an outward-pushing perturbation. A short self-timed
213 break was given every 30 trials in all experiments.

214

215 After they completed the experiment, participants in all four experiments were given a brief
216 questionnaire to assess their awareness of the perturbation as well as whether they adopted any
217 strategy and, if so, what that strategy was.

218

219 *Quantification and statistical analysis*

220 F1 and F2 were tracked offline using `wave_viewer` (Niziolek & Houde, 2015), which provides a
221 MATLAB GUI interface to formant tracking using Praat (Boersma & Weenink, 2019). Linear
222 predictive coding (LPC) order and pre-emphasis values were adjusted individually for each
223 participant. All trials were first checked manually for errors in production (e.g., if the participant
224 said the wrong word). Vowel onset and offset were detected automatically using a participant-
225 specific amplitude threshold and errors in the location of these automatically-defined landmarks
226 were manually corrected using the audio waveform and spectrogram. Vowel onset was marked at
227 the point where periodicity was visible in the waveform and formants were visible in the
228 spectrogram. Vowel offset was marked at the point where formants, particularly F2 and higher,
229 were no longer visible. Errors in formant tracking were manually corrected by adjusting the LPC
230 order or pre-emphasis value on a trial-specific basis. In total, 1.9% of the data were excluded due
231 to production errors or unresolvable errors in formant tracking. For each trial, F1 and F2 values
232 were calculated by averaging formants from a 50-ms segment at both the beginning (vowel
233 onset) and the middle (vowel midpoint) of each vowel.

234

235 The primary goal of the analysis was to test how variability changed across the different phases
236 of each experiment. For offline analysis, the exposure phase was equally divided into early
237 exposure and late exposure phases to make sure each phase contained a similar number of trials.
238 Variability within each experimental phase was calculated as the average of the 2D distances in
239 F1/F2 space between each production of a vowel and the center of the distribution for that vowel
240 in that phase, measured from the first 50 ms of vowel. In order to test how variability may
241 change in specific dimensions, we additionally calculated formant variability separately along
242 the F1 and F2 axes, as well as along the major and minor axes of produced variability in the
243 baseline phase. The variability along the F1 and F2 axes was defined as the standard deviation
244 (SD) of F1 and F2 values of all productions of a stimulus word during each phase. For the major
245 and minor axes of variability, an ellipse which represents a 95% confidence interval of trials in
246 F1-F2 spaces was fitted for each stimulus word and each experimental phase using the Principal
247 Components method. The vector representing the F1 and F2 values for each trial was projected
248 into a component along the major axis of the fitted ellipse and a component along the minor axis
249 perpendicular to the major axis. The variability along the major and minor axes was defined as

250 the standard deviation (SD) of projected values along that axis of all productions of a stimulus
251 word during each phase.

252

253 We additionally measured vowel centering, a measure of within-trial correction for variability,
254 calculated by the change in variability from vowel onset (first 50 ms) to vowel midpoint (middle
255 50 ms). Vowel centering allows us to determine whether participants altered their within-trial
256 control of variability in response to the perturbation (Niziolek & Guenther, 2013; Niziolek &
257 Kiran, 2018; Niziolek & Parrell, 2021). Centering was measured separately for each vowel in
258 each experimental phase.

259

260 Repeated measure analyses of variance (ANOVAs) were conducted separately for the variability
261 and centering results and for each experiment, with phase and vowel identity as within-subject
262 factors. For the control data and Experiments 1-3, data from baseline, late exposure and washout
263 phases were included in the repeated ANOVAs, while data from baseline, inward-pushing
264 exposure, and outward-pushing exposure phases were included in Experiment 4. Post-hoc
265 comparisons (paired t-tests) were only conducted in the event of a significant main effect of
266 phase or interaction. As an exploratory analysis, multiple regression was conducted in each
267 experiment to determine whether produced variability changes could be predicted by baseline
268 variability and vowel identity. Finally, variability changes along F1/F2 or major/minor axes were
269 examined separately by three-way repeated measures ANOVAs which included phase, vowel
270 identity and axis (i.e. F1/F2 or major/minor) as within-subject factors. The significance level for
271 all statistical tests was $p < .05$, with a Bonferroni correction for multiple comparisons for post-
272 hoc tests.

273 All statistical analyses were conducted in R (R Core Team, 2019). Repeated measures ANOVAs
274 and pairwise paired t-tests were conducted with the `rstatix` package (Kassambara, 2021), in
275 which partial eta squared (η_p^2) and Cohen's d were calculated for repeated ANOVAs and paired
276 t-tests, respectively, to determine effect size for statistically significant effects. Greenhouse-
277 Geisser Correction was applied automatically to correct the degrees of freedom when sphericity
278 was violated. Multiple regression models were constructed using the package `stats`. Data and
279 associated code is available at <https://osf.io/stjc9/>. Some functions rely on additional code
280 available at <https://github.com/carrien/free-speech>.

281

282 *Model simulations*

283 In order to assess the potential mechanisms underlying the patterns of variability changes
284 observed in Experiments 1-4, we conducted a simulation of speech behavior using a version of
285 the well-established state space models that have been used in studies of sensorimotor adaptation
286 to sensory perturbations in both limb (Baddeley et al., 2003; Donchin et al., 2003; Thoroughman

287 & Shadmehr, 2000) and speech (Daliri, 2021; Daliri & Dittman, 2019) motor control.
288 Importantly, analogous models have also been used successfully in studies of sensorimotor
289 corrections for self-produced variability in reaching (Ahn et al., 2016; Blustein et al., 2021;
290 Scheidt et al., 2001), similar to the current experiments. The model is:

$$\begin{aligned} 291 & \\ 292 & m_t = Am_{t-1} + Be_{t-1} \\ 293 & x_t = m_t + N(0, K) \\ 294 & e_t = -Gx_t \\ 295 & \end{aligned}$$

296 This model assumes that the intended production (m) on a given trial (t) aims to achieve a
297 particular target (by convention, set to 0) based on a weighted contribution of the production on
298 the previous trial ($t - 1$) and the error (e) experienced on the previous trial. A is the forgetting
299 factor that determines the contribution of the previous trial, and B represents the sensitivity of the
300 system to errors. The error e is the difference between the actual production outcome (x) and the
301 target, where the production x is the result of the intended production m plus noise (N) drawn
302 from a gaussian distribution with a mean of 0 and a standard deviation of K . In order to account
303 for the gains applied to the auditory feedback in this experiment, the observed outcome x is
304 multiplied by a gain factor G to derive the error. Aside from the addition of the gain factor G , this
305 is identical to the formulation in Ahn et al., (2016), if $A = 1$.

306
307 While the state space model is often used to model sensorimotor adaptation to external
308 perturbations, the addition of the noise term on the final motor output also permits modelling of
309 correction for self-produced variability (Ahn et al., 2016; Blustein et al., 2021; Scheidt et al.,
310 2001). For simplicity, we initially set A to 1 (Ahn et al., 2016) and B to 0.1 given previous results
311 fitting state space models to adaptation in speech (Daliri, 2021; Daliri & Dittman, 2019). We
312 then estimated an initial value of K that would generate an observed distribution similar to the
313 experimentally observed variability in the baseline phase across experiments (roughly 30 mels).
314 Because the errors in this model are generated stochastically, we need a large number of
315 simulations to achieve an accurate estimate of the underlying distribution of variability that
316 would be generated through this model. To that end, we ran batches of 1000 simulations while
317 varying K . Each simulation consisted of 40 trials, equivalent to the number of trials used to
318 estimate variability in our experimental data. Our initial simulations indicated that setting K to 30
319 resulted in a mean observed variability close to the observed values in the baseline phase across
320 experiments. Varying A within the range reported in (Daliri, 2021) had very minor effects on the
321 results; the effects of varying B are explored below.

322
323 Our goal in modeling was to assess the potential causes of the increase in variability observed in
324 the behavioral data in Experiments 1-4. To do this, we systematically varied the underlying
325 variability K as well as the sensitivity to errors B . K varied from 30 to 40 in steps of 0.5. B varied
326 from 0 to 1 in steps of 0.05, where 0 would represent no correction for observed errors and 1

327 would represent full correction. For each step of K and B , 1000 simulations (40 trials each) were
328 conducted. The standard deviation of x was calculated for each simulation, and the mean of these
329 standard deviations was calculated to generate an estimate of the expected variability with that
330 particular parameter set. We ran separate simulations for the inward-pushing perturbation in
331 Experiment 1 (setting G to 0.5) and the outward-pushing perturbation in Experiment 2 (setting G
332 to 1.75).
333

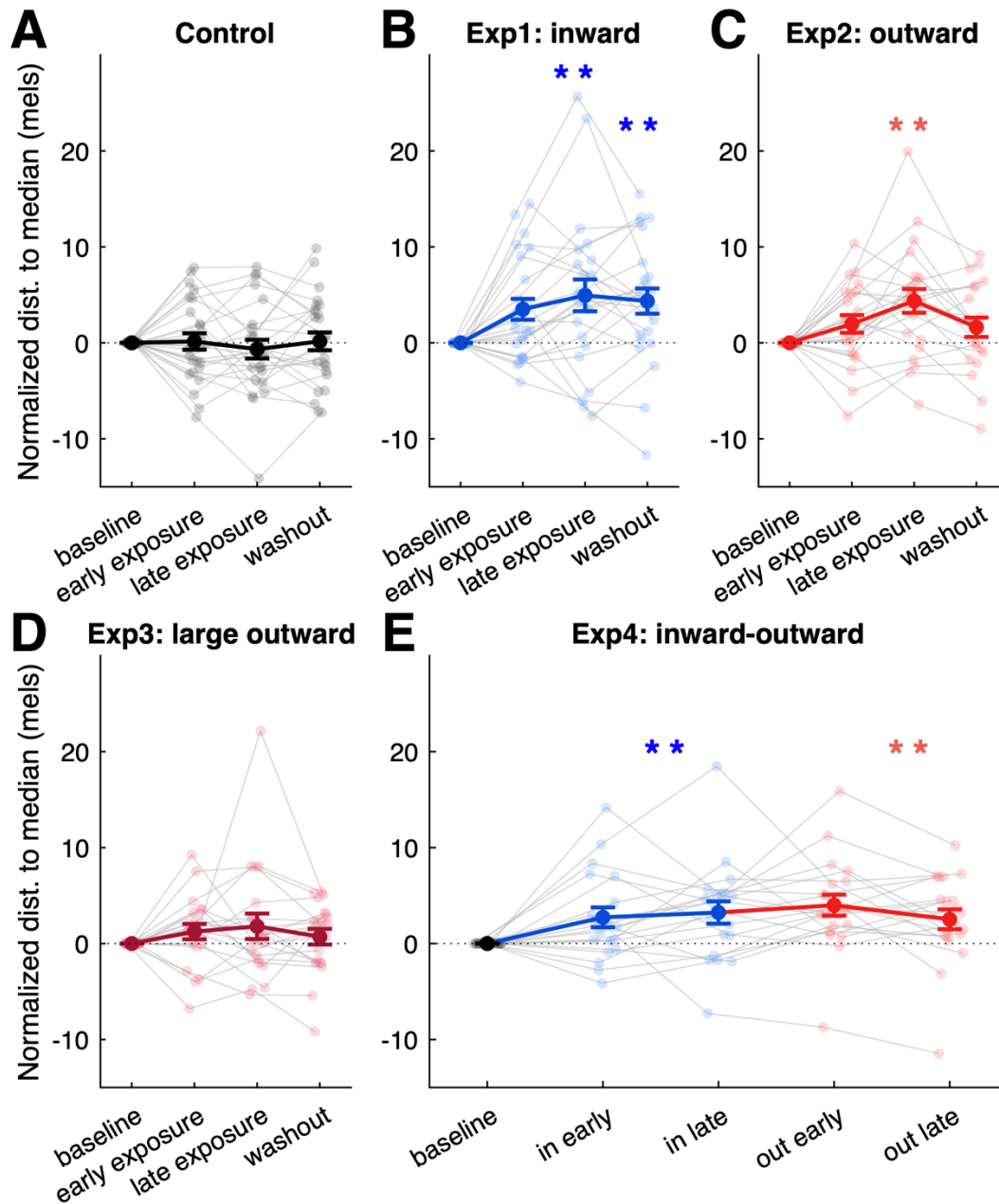
334 **Results**

335 *Overall variability changes*

336 Over four experiments, we implemented an auditory perturbation designed to increase or
337 decrease participants' perceived trial-to-trial variability without affecting their overall mean. To
338 confirm that this was achieved, the perturbations applied to F1 or F2 frequencies during the
339 exposure phase in each experiment were averaged for each participant and compared against
340 zero (no mean formant value change) using one-sample t-tests. The mean F1 perturbations
341 ranged between 2.5 and 13.4 mels in Experiments 1-4 and significantly differed from zero only
342 in Experiment 3 (large outward-pushing, mean F1 perturbation: 13.4 mels, $t(20) = 2.47$, $p =$
343 0.023). Similarly, the mean F2 perturbations ranged between -1.23 and 12.39 mels in
344 Experiments 1-4, with no significant difference from zero in any case (uncorrected $p > 0.05$). It
345 is worth mentioning that the just noticeable difference (JND) in F1 and F2 for isolated English
346 vowels is around 14 mels and 20 mels, respectively (Kewley-Port & Watson, 1994).
347

348 We then measured how participants changed their produced variability after the perceived
349 variability had been increased or decreased. As a control, we analyzed an existing dataset (Parrell
350 & Niziolek, 2021) with no auditory perturbation. As expected, the control group showed no
351 change in variability over the course of the experiment (Figure 2A, $F(2,48) = 0.434$, $p = 0.650$).
352

353 Participants exposed to the **inward-pushing perturbation (Experiment 1)** increased produced
354 variability (Figure 2B, main effect of phase: $F(2,46) = 7.55$, $p < 0.001$, $\eta_p^2 = 0.25$) while the
355 perturbation was applied (+5.0 mels, $t(71) = 4.59$, $p < 0.001$, $d = 0.54$) as well as after it was
356 removed (+4.4 mels, $t(71) = 4.89$, $p < 0.001$, $d = 0.58$). Unexpectedly, participants exposed to
357 the **outward-pushing perturbation (Experiment 2)** also increased produced variability (Figure
358 2C, main effect of phase: $F(2,40) = 7.93$, $p < 0.001$, $\eta_p^2 = 0.28$). However, unlike the maintained
359 variability change seen in the inward-pushing perturbation, the increased variability during
360 exposure (+4.3 mels, $t(65) = 4.10$, $p < 0.001$, $d = 0.50$) returned to near-baseline levels when the
361 outward-pushing perturbation was removed (+1.6 mels, $t(62) = 1.93$, $p = 0.175$).



362
363

364 **Figure 2. Baseline-normalized variability changes.** Individual (small transparent dots, thin
365 lines) and group means (large solid dots, thick lines) of baseline-normalized variability
366 (normalized by subtracting the average value in the baseline from the remaining trials) in the
367 baseline, exposure and washout phases. Error bars show standard error. ** indicates significant
368 change ($p < 0.001$) from baseline.

369
370

371 We reasoned that the failure to find the hypothesized reduction in variability in response to the
372 outward-pushing field could potentially be attributed to the size of the perturbation used.

373 Although this perturbation was slightly larger than in Experiment 1 (75% vs 50% of the distance
374 to vowel target), it may still have been too small to induce participants to reduce their variability.
375 **Experiment 3 (large outward-pushing perturbation)** aimed to delineate this by testing another
376 group of participants who received outward-pushing perturbations with a 200% increase in
377 distance to the center of vowel distribution (vs 75% in Experiment 2). However, participants
378 exposed to this large outward-pushing perturbation did not change their produced variability
379 (Figure 2D, main effect of phase: $F(1.55,30.95) = 1.39, p = 0.26$).

380
381 Together, the results of both experiments employing outward-pushing perturbation fields
382 (Experiment 2 and 3) suggest these perturbations do not drive participants to produce
383 compensatory reductions in variability. One possibility for this behavior is that speech
384 movements are already produced at or near the lower limit of an individual speaker's precision
385 ability. **Experiment 4 (inward-outward pushing perturbation)** aimed to test this possibility by
386 examining whether an outward-pushing perturbation can reduce produced variability of
387 participants *after* their variability has increased above baseline levels due to exposure to an
388 inward-pushing perturbation as seen in Experiment 1. Participants significantly changed their
389 produced variability during the course of Experiment 4, reflected by a main effect of phase
390 (Figure 2E, $F(2,38) = 9.71, p < 0.001, \eta_p^2 = 0.34$). As expected given the results of Experiment 1,
391 variability increased when the inward-pushing perturbation was applied (+4.1 mels, $t(59) = 4.77,$
392 $p < 0.001, d = 0.62$). However, participants did not change their produced variability back to the
393 baseline when receiving the outward-pushing perturbation in the following phase (+4.1 mels,
394 $t(59) = 4.13, p < 0.001, d = 0.53$). While this result replicates the increase in variability observed
395 during the exposure phase in Experiments 1 and 2, it suggests the failure to find the expected
396 reduction in variability in Experiments 2 and 3 was not caused solely by a “lower limit” on
397 variability.

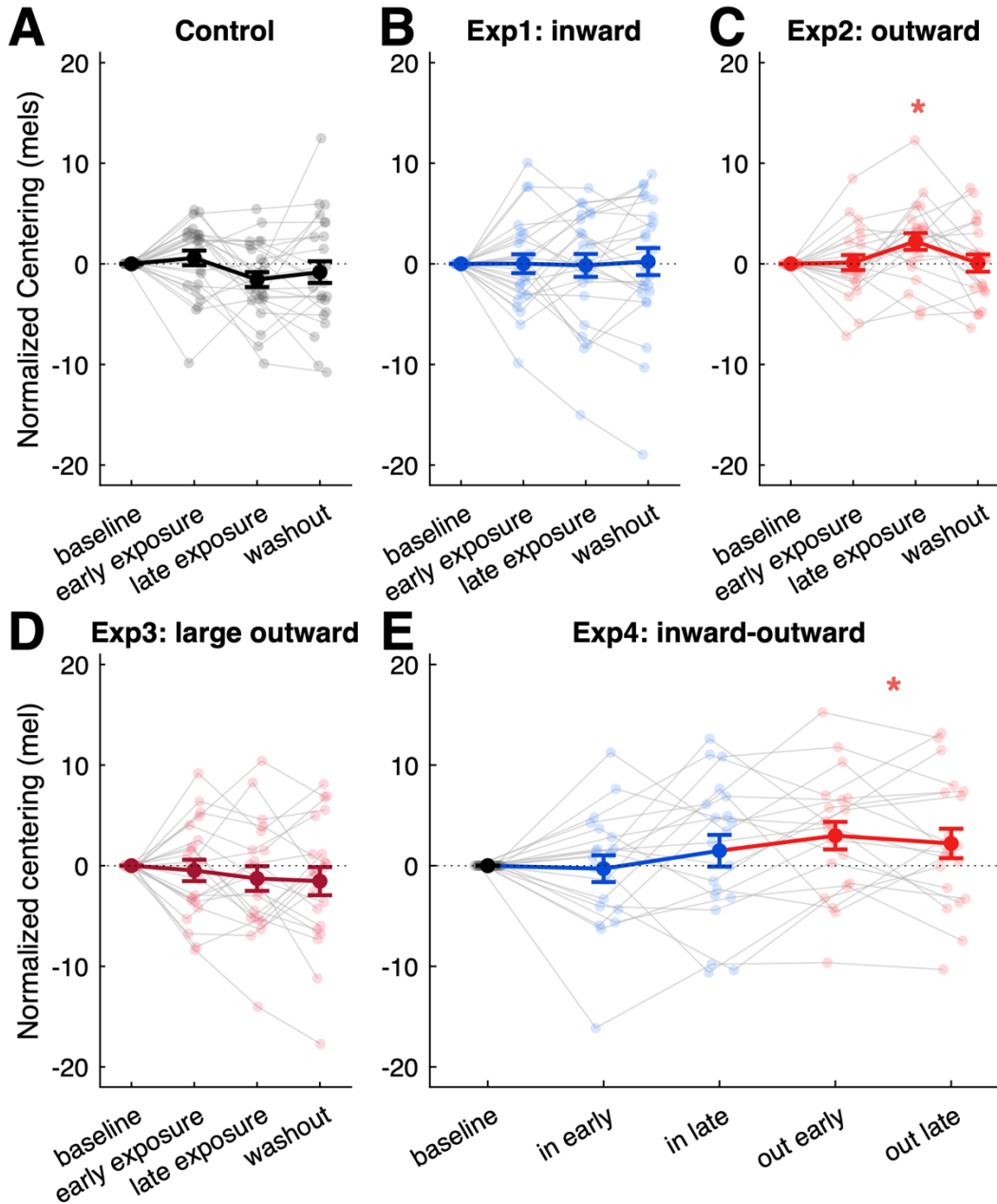
398
399 Finally, while not the primary focus of the study, we evaluated whether participants adjusted
400 their variability differently along the major/minor axes of variation using three-way repeated
401 ANOVAs including three within-subject factors (phase, vowel, axis). This analysis replicated the
402 results of overall variability changes (distances in F1-F2 spaces): a significant main effect of
403 phase was found in Experiments 1, 2, 4, but not in the control data or in Experiment 3. Perhaps
404 more interestingly, no significant two-way interaction between phase and measure was found in
405 any of these experiments, suggesting participants adjusted their variability along the major- and
406 minor-axis similarly. Similar results were found from models of F1 and F2 variability. See
407 Supplemental Material, Figures S2, S3 and Tables S1, S2 for detailed results and statistics.

408

409 ***Centering changes***

410 In order to determine whether participants adjusted their within-trial control of variability in
411 response to the perturbation, we additionally measured vowel centering (Niziolek & Guenther,

412 2013; Niziolek & Kiran, 2018; Niziolek & Parrell, 2021), the reduction in variability from vowel
413 onset (first 50 ms) to vowel midpoint (middle 50 ms). Similar to the analyses on variability, the
414 control group did not show any change in centering over the course of the experiment (Figure
415 3A, $F(2,48) = 1.45, p = 0.244$). However, unlike the overall variability changes observed above,
416 no change in centering was seen in participants who received the **inward-pushing perturbation**
417 (Experiment 1, Figure 3B, $F(1.49, 34.27) = 0.06, p = 0.896$). In contrast, centering did increase
418 when participants were exposed to the **outward-pushing perturbation** (Experiment 2, Figure
419 3C, main effect of phase: $F(2,40) = 3.94, p = 0.027, \eta_p^2 = 0.165$), suggesting these participants
420 became more responsive to errors. This increase in centering during exposure (+ 2.2 mels, $t(65)$
421 $= 3.33, p = 0.004, d = 0.41$) was not retained during the washout phase (+ 0.1 mels, $t(62) = 0.11,$
422 $p = 1.000$). Contrary to the increase in centering observed in Experiment 2, participants who
423 received a **large outward-pushing perturbation** in Experiment 3 (i.e. 200% increase in distance
424 to the vowel target) did not show an increase in centering over the course of experiment (Figure
425 3D, $F(2,40) = 0.87, p = 0.428$), actually tending to slightly decrease centering during and after
426 exposure. In Experiment 4 (**inward-outward pushing perturbation**), participants showed
427 changes in centering over the course of the experiment, indicated by a marginally significant
428 main effect of phase (Figure 3E, $F(1.4, 26.5) = 3.15, p = 0.075, \eta_p^2 = 0.142$). More specifically,
429 the results replicated the patterns observed in Experiments 1 and 2: the initial inward-pushing
430 perturbation did not induce any significant increase in centering (+ 0.53 mels, $t(59) = 0.38, p = 1.$
431 000), while the subsequent outward-pushing perturbation did (+ 2.7 mels, $t(59) = 2.27, p = 0.05,$
432 $d = 0.29$).



433

434

435 **Figure 3. Baseline-normalized centering changes.** Individual (small transparent dots, thin
436 lines) and group means (large solid dots, thick lines) of baseline-normalized centering
437 (normalized by subtracting the average value in the baseline from the remaining trials) in the
438 baseline, exposure and washout phases. Error bars show standard error. * indicates significant
439 change ($p < 0.05$) from baseline.

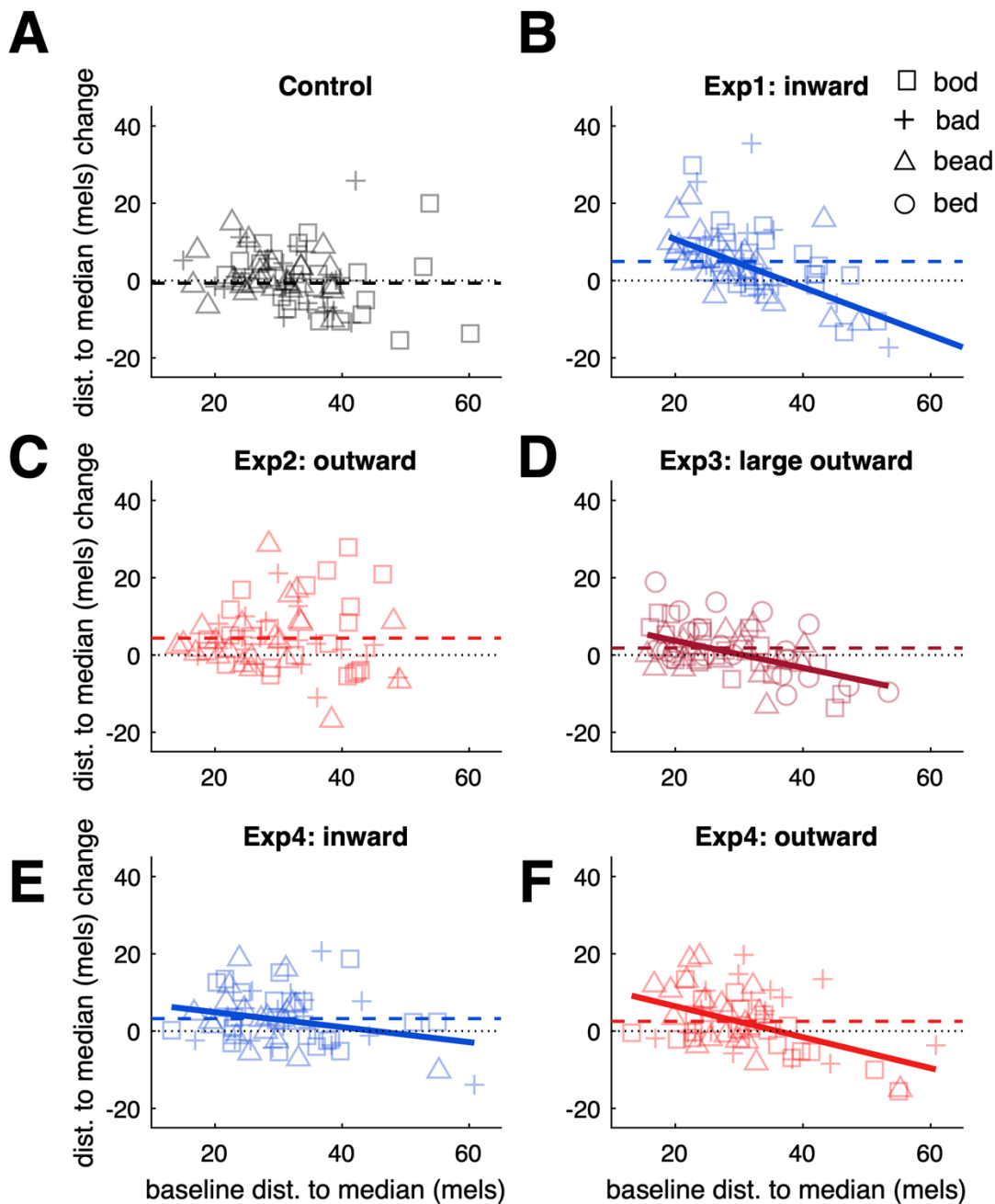
440

441 ***Correlation between variability change and baseline variability***

442 Multiple regression was conducted in each experiment to determine whether changes in
443 produced variability could be predicted by baseline variability and vowel identity at the
444 individual level (see Figure 4). Results showed that baseline variability was not predictive of the
445 individual change in variability observed during the late exposure phase in either the control ($\beta =$
446 $-0.18, p = 0.094$) or Experiment 2 (**outward-pushing perturbation**, $\beta = -0.024, p = 0.846$).
447 However, this relationship was observed in the Experiment 1 (**inward-pushing perturbation**, β
448 $= -0.44, p < 0.001$), such that participants with lower variability in the baseline phase showed
449 larger variability increases. In Experiment 3 (**large outward-pushing perturbation**), where
450 there was no consistent change in variability over the course of the experiment, baseline
451 variability was nonetheless predictive of individual changes in variability ($\beta = -0.38, p = 0.002$):
452 participants with higher variability in the baseline phase tended to decrease variability, and vice
453 versa. Experiment 4 (**inward-outward pushing perturbation**) replicated the result from
454 Experiment 1: a significant correlation was observed between baseline variability and variability
455 change induced by inward-pushing perturbation ($\beta = -0.19, p = 0.049$). Perhaps surprisingly
456 given the results of Experiment 2, a similar relationship between baseline variability and
457 variability change was also observed during later outward-pushing perturbation ($\beta = -0.37, p <$
458 0.001). However, it should be noted that there was a highly significant correlation in variability
459 between the two perturbation phases ($\beta = 0.54, p < 0.001$), suggesting that the correlation
460 observed during the outward-pushing perturbation is likely a carry-over effect of the inward-
461 pushing perturbation which happened between baseline and outward-pushing phases.

462

463



464

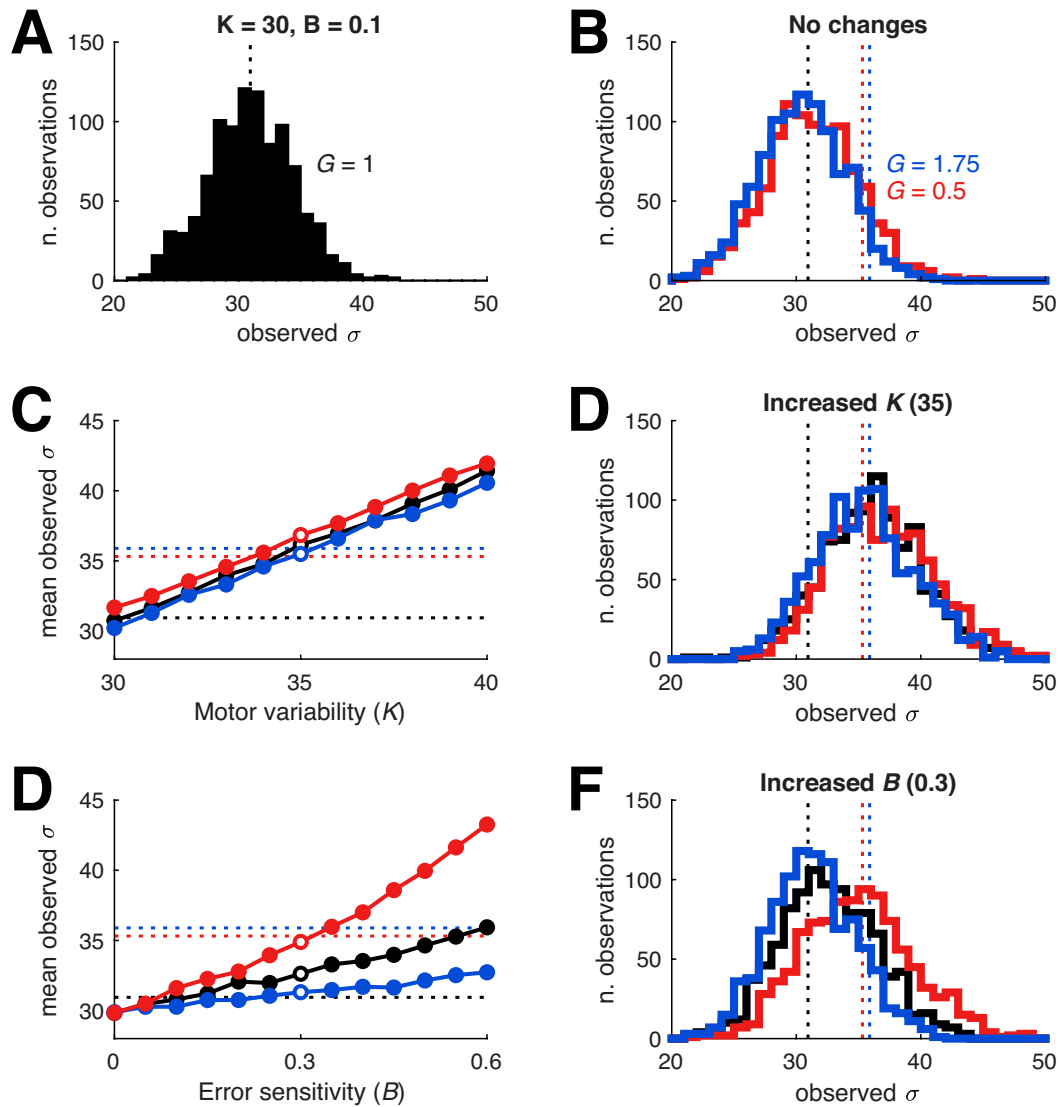
465

466 **Figure 4. Correlation between average baseline variability and variability change across**
467 **individuals.** Each data point represents the average production of one stimulus word, indicated
468 by different markers. Each participant thus contributed three data points. Note that the stimulus
469 words were “bead”, “bad”, and “bed” in Experiment 3, and “bead”, “bad”, and “bod” in the
470 other experiments. Significant correlations are indicated by solid lines showing the least-squares
471 fit to the data points. The group-averaged (baseline-normalized) variability changes are
472 indicated by colored dashed lines.

473 ***State space simulations***

474 Using a modified version of the typical state space model that has been shown to account for
475 observed motor variability in reaching, we systematically varied the magnitude of motor
476 variability (K) and the sensitivity of the system to observed errors (B). Results of these
477 simulations are shown in Figure 5. We found that, predictably, increases in the underlying motor
478 variability increased the observed variability in motor production for both the inward- and
479 outward-pushing perturbations (Figure 5C). Values of K near 35 produced observed variability
480 measures consistent with those in both Experiments 1 and 2 (Figure 5D). Conversely, we found
481 that changes in error sensitivity had differential effects on motor variability depending on the
482 perturbation field. For the inward-pushing perturbation, variation in error sensitivity had very
483 minor effects on motor output. Conversely, for the outward-pushing perturbation, increases in
484 error sensitivity led to large increases in motor variability (Figure 5E). Values of B near 0.3
485 resulted in a good match for the observed variability in Experiment 2, while no values of B
486 provided a good match for Experiment 1 (Figure 5F). In the latter case, the maximum increase in
487 observed variability over baseline was 3.7 where $B = 1$, still substantially less than the
488 experimentally observed value. These results indicate that the observed increase in Experiment 1
489 can only be explained by an increase in the underlying motor variability, while the increase
490 observed in Experiment 2 could arise through either an increase in controlled variability or an
491 increase in the sensitivity to auditory errors.

492



493
 494 **Figure 5. State space simulations of motor variability.** For all panels, the dashed black line
 495 represents the mean variability with model parameters set to match participant behavior in the
 496 baseline phase, the dashed blue line represents the increase in variability observed in
 497 Experiment 1 (4.9 mels above baseline), and the dashed red line represents the increase in
 498 variability observed in Experiment 2 (4.4 mels above baseline). (A) shows distributions of
 499 observed variability for 1000 simulations where the gain applied to the observed error is 1,
 500 simulating the baseline phase in both Experiments 1 and 2. (B) shows similar distributions from
 501 simulations with gains set to 0.5 (blue) and 1.75 (red), mirroring the perturbations applied in
 502 Experiment 1 and Experiment 2, respectively, with no changes to other model parameters. (C)
 503 shows the result of systematically varying underlying motor variability K on observed variability
 504 with gains set to 1 (black), 0.5 (blue), and 1.75 (red). The open circles in the left panel represent
 505 the value for K used in the simulations shown in (D). (D) shows the distribution of observed
 506 variability with K set to 35, chosen to roughly match the increase in variability observed in

507 *Experiments 1 and 2. (E) as for (C), showing the effect of varying the error sensitivity parameter*
508 *B. Values of B greater than 0.6 (not shown) resulted in values observed variability > 50 where G*
509 *= 1.7 and < 34 where G = 0.5. (F) shows the distribution observed with B set to 0.3, chosen to*
510 *match the increase in variability observed in Experiment 2.*
511

512 *Awareness of perturbation*

513 Most of the participants in Experiment 1, 2, and 4 (41/66) were not aware of the perturbation
514 applied to their vowels. Although 25/66 did indicate that they thought their speech was somehow
515 manipulated, only 1 participant correctly identified it as a change to their vowels/consonants. In
516 contrast, nearly all participants (19/21) in Experiment 3 (large outward-pushing) reported that
517 they thought they received a perturbation, and almost half (9/21) correctly identified it as a
518 manipulation to their vowels. No participants reported a strategy that aimed to address the
519 applied perturbation in any of the four experiments.
520

521 **Discussion**

522 In a set of four experiments, we examined whether variability can be actively regulated in a
523 complex, well-practiced motor task: speech production. Specifically, we examined whether
524 participants would adjust their produced variability when they were exposed to real-time
525 auditory perturbations that increased or decreased their perceived variability. Our results showed
526 that introducing a perturbation that reduces perceived variability (Experiment 1: inward-pushing)
527 leads to increases in produced vowel variability that remain even when normal feedback is
528 restored, suggesting that variability is monitored and regulated over relatively long time scales.
529 Perhaps surprisingly, a perturbation that increases perceived variability (Experiment 2: outward-
530 pushing) also increased produced variability, though such variability change was not maintained
531 when the perturbation was removed. In Experiment 3, participants who received a large outward-
532 pushing perturbation (i.e. a 200% increase in distance to the vowel target) did not change their
533 produced variability over the course of the experiment. This rules out the possibility that the
534 failure to find the hypothesized reduction in variability in response to the outward-pushing
535 perturbation (Experiment 2) is due to inadequate perturbation magnitude. Finally, we found that
536 an outward-pushing perturbation cannot reduce participants' produced variability even *after* their
537 variability has increased above baseline levels as a result of the inward-pushing perturbation
538 (Experiment 4: inward-outward pushing).
539

540 To our knowledge, this is the first study showing that variability in speech production can be
541 actively controlled, consistent with recent theories that have highlighted the adaptive value of
542 motor variability in other motor domains (Davids et al., 2003; Herzfeld & Shadmehr, 2014;
543 Tumer & Brainard, 2007). From this perspective, motor variability can be actively generated,

544 regulated and used by the brain to improve motor performance, reduce costs and explore new
545 solutions (Shadmehr et al., 2016). Early evidence shows that skilled performers are able to
546 upregulate the level of motor variability in each joint of the upper arm to meet the change of task
547 demands/constraints, while less skilled performers, in comparison, tend to have rigidly fixed
548 motor variability that is not fine-tuned with task constraints (Arutyunyan et al., 1968, see Newell
549 & Vaillancourt, 2001 for a review). More recent work using computational models also found
550 that force variability and the resulting kinematic variability are not generated primarily by
551 random “motor noise”, and emphasize the importance of other sources of force variability which
552 can be tuned as needed by distributed sensorimotor systems (Nagamori et al., 2021). The results
553 from the current study extend previous work and provide support for this perspective by showing
554 the motor system closely monitors sensory variability and uses such information to actively
555 regulate the motor variability, even in complex and well-practiced behaviors such as natural
556 speech.

557
558 Surprisingly, participants exposed to both inward-pushing and outward-pushing perturbations
559 increased their produced variability. These results could be interpreted as a general variability
560 increase induced by either repetitive productions of many utterances or a non-specific auditory
561 perturbation (that is, any kind of auditory perturbation would lead to an increase in formant
562 variability). However, the results from our data as well as previous studies indicate that both of
563 the possibilities are unlikely. First, in analyzing an existing dataset where participants produced
564 460 utterances with normal auditory feedback (Parrell & Niziolek, 2021), we found that speakers
565 did not significantly change their produced variability. This confirms that under normal
566 circumstances speakers produce vowels with a level of variability that is relatively stable over
567 time. Second, previous work has shown that participants do not change their formant variability
568 in response to consistent auditory perturbation of both F1 and F2 (i.e. shifts of 240 Hz in F1 and
569 300 Hz in F2) (Nault & Munhall, 2020). Results from Experiment 3 are consistent with this
570 result, as participants who received a large outward-pushing perturbation did not exhibit any
571 significant change in their variability. These results, taken together, rule out the possibility of a
572 general non-specific increase in variability induced by auditory perturbations.

573
574 In our experimental data, the increased variability induced by the auditory perturbation was
575 maintained even when normal feedback was restored in Experiment 1, but not in Experiment 2.
576 This difference suggests different mechanisms may have led to the increased variability observed
577 in the two experiments. As an attempt to disentangle these potential mechanisms, we conducted a
578 simulation using a state space model of error correction. Model simulations identified two
579 distinct mechanisms that could lead to the observed increase in produced variability: an increase
580 in controlled variability or an increase in the sensitivity to auditory errors. More specifically,
581 these simulations indicated that the observed increase in Experiment 1 (inward-pushing) can only
582 be explained by an increase in controlled variability, while the increase observed in Experiment
583 2 (outward-pushing) could arise through either of these two mechanisms. Importantly for

584 Experiment 2, a return to unperturbed auditory feedback in the washout phase would not
585 necessarily cause an immediate decrease in controlled variability, but would directly lead to a
586 change in variability related to trial-to-trial correction for errors even with a constant error
587 sensitivity (see Figure 5). These modeling results suggest the increase in observed variability in
588 response to the inward-pushing perturbation in Experiment 1 was likely driven by a relaxation of
589 controlled variability, as the perturbation “frees” the motor system to be more variable without
590 any loss in perceived accuracy. Conversely, the results from the Experiment 2 are most
591 consistent with an increase in error sensitivity caused by the outward-pushing field rather than a
592 change in controlled variability. Consistent with this difference, we found a significant
593 correlation between produced variability changes and baseline variability in Experiment 1 but
594 not in Experiment 2. One possible explanation for the pattern observed in Experiment 1 is that
595 participants who are naturally more variable may take advantage of the natural consequences of
596 the inward-pushing perturbation to reduce overall perceived variability, while less variable
597 individuals may relax their (presumably stricter) regulation of variability without negatively
598 affecting their perceived variability. In contrast, in Experiment 2, we would not expect that
599 changes in error sensitivity would be related to levels of baseline variability. In brief, the
600 combined results of behavioral and model simulation point to the two distinct mechanisms that
601 may have led to the observed increase in variability in response to inward-pushing and outward-
602 pushing perturbation.

603
604 It is worth mentioning that while it would be ideal to compare the predictions generated by the
605 state space model to empirical estimates of error sensitivity, limitations of the experimental
606 design prevent us from being able to estimate such parameters. Even with recent advances in this
607 regard (Blustein et al., 2021), accurate estimation of these parameters requires 100-200
608 sequential trials. In all experiments, participants produced 40-50 repetitions of each word per
609 phase, in a pseudo-randomized order. Thus, our protocol provides both an insufficient number of
610 trials as well as a potential confound of producing multiple intervening movements between
611 repetitions of the same item. Future work should explore whether current methods that are
612 known to provide accurate estimates of error sensitivity in reaching could be applied to speech
613 motor control with a more appropriate experimental design.

614
615 We additionally measured vowel centering, a measure of within-trial correction for variability
616 (Niziolek & Guenther, 2013; Niziolek & Kiran, 2018; Niziolek & Parrell, 2021). Unlike the
617 overall variability changes observed in both kinds of perturbation, an increase in centering was
618 only seen in participants who received the outward-pushing perturbation in Experiment 2 and in
619 the outward-pushing phase of Experiment 4. This suggests that participants exposed to outward-
620 pushing perturbations became more responsive to within-trial errors (i.e. larger within-trial
621 feedback gains). Although it is not clear that such an increased sensitivity to errors for within-
622 movement corrections is directly related to the error sensitivity related to cross-trial changes (i.e.
623 gains on trial to trial learning), it is possible error sensitivity is shared across both processes. This

624 would be consistent with the fact that we see increased centering only in Experiment 2, precisely
625 where modeling results suggest an increase in trial-to-trial error sensitivity.

626

627 The only case where we did not see an increase in centering or variability was Experiment 3
628 (large outward-pushing perturbation). It is possible that the larger perturbations used in this
629 experiment were discounted by the sensorimotor system, such that they had limited effects on
630 speech production. This is consistent with previous studies that have similarly shown attenuated
631 responses to large auditory feedback perturbations (Burnett et al., 1998; Scheerer et al., 2013),
632 potentially because large perturbations are treated as externally induced, rather than self-
633 produced errors (Korzyukov et al., 2017). This is also in line with the perturbation awareness
634 results: almost half of the participants in Experiment 3 correctly recognized the perturbation
635 applied on vowels, while only one participant did across all other experiments.

636

637 Together, the results of Experiments 2-4 suggest that a sensory perturbation that increases
638 perceived variability does not drive participants to produce compensatory reductions in speech
639 variability, contrary to previous results in non-speech motor control which showed task-relevant
640 variability can be reduced when needed or after repeated practice (Kang et al., 2004; van Beers et
641 al., 2013; Wong et al., 2009). One potential explanation for our failure to observe the
642 hypothesized decrease in variability in response to the outward-pushing perturbation in
643 Experiment 2 is that speech is already produced at the lower limits of possible task-related
644 variability, consistent with the predictions of the Uncontrolled Manifold hypothesis and Optimal
645 Feedback Control (Harris & Wolpert, 1998; Scholz & Schöner, 1999; Todorov, 2004). This
646 could also explain the lack of change in produced variability when the outward-pushing
647 perturbation occurred after variability had increased above baseline levels (Experiment 4): it is
648 possible that controlled variability here returned to baseline levels (without reducing past that
649 point), but was counteracted by the increase of error sensitivity induced by outward-pushing
650 perturbations.

651

652 Another possibility is suggested by the fact that in reaching tasks, only task-relevant variability
653 has been observed to reduce experimental tasks, while task-irrelevant variability remained high
654 (Robertson & Miall, 1997; Scholz & Schöner, 1999). Thus, it is possible that participants in the
655 current experiments reduced variability along a particular dimension, even if their overall
656 variability increased. To explore this possibility, we calculated changes in variability along
657 participant-specific major and minor axes of variability in the baseline phase, assuming that the
658 minor axis of variability may be more tightly controlled (Scholz & Schöner, 1999). However, we
659 found that participants produced similar changes in variability along the major- and minor-axis,
660 suggesting speech motor system might control variability more globally compared to the
661 selective regulation in certain components of movement variability observed in non-speech
662 motor control (Abe & Sternad, 2013; Sternad et al., 2011). It is important to note, however, that
663 the major- and minor-axis in speech production are not necessarily equivalent to the task-

664 relevant and task-irrelevant dimensions in non-speech control, and indeed may both contribute to
665 task performance. Further work is needed to clarify this point.

666

667 In summary, we have shown that individuals modify their produced variability when their
668 perceived trial-to-trial variability is altered. Decreases in perceived variability lead to increases
669 in produced variability, likely due to loosened restrictions on variability production, particularly
670 in individuals with inherently low variability. These changes are retained even when the
671 perturbation is removed, suggesting that the monitoring and regulation of variability acts
672 relatively slowly. Conversely, variability also increases in response to perturbations which
673 increase perceived variability, potentially due to increases in error sensitivity as participants try
674 to correct for the perceived errors. Together, these results are consistent with recent evidence that
675 suggests motor variability should be viewed as an important feature of how the sensorimotor
676 system operates and learns rather than as the inevitable and unintended consequence of motor
677 noise. Our results also highlight the importance of having a better understanding of motor
678 variability during speech production, which has been largely overlooked in current theories and
679 models of speech motor control.

680

681

682

683 **Acknowledgments**

684 This work was supported by NIH grant R01 DC019134 and a grant awarded through the
685 University of Wisconsin–Madison Fall Research Competition.

686

687

688

689

690

691

692

693

694

695

696

697

698

699

700

701

702

703

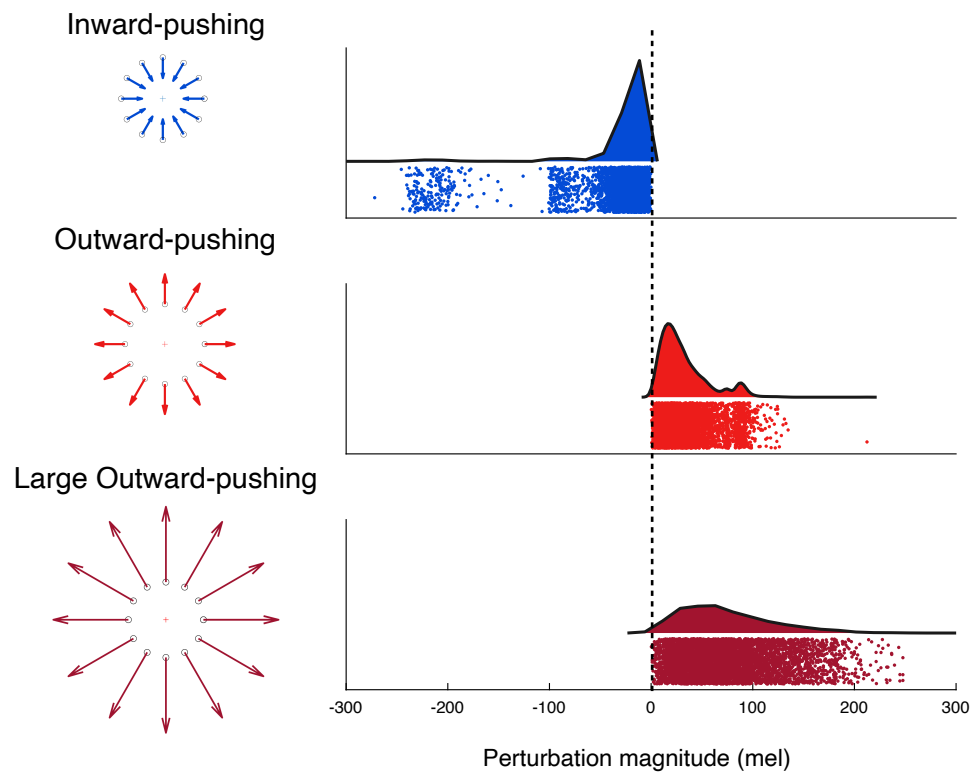
704

705

706

707
708
709
710
711
712
713
714
715

Supplementary materials



716
717
718
719
720
721
722
723
724

Figure S1. Schematic (left) and raincloud plot (right) of perturbation applied to speech vowel formants: inward-pushing pushing (Exp 1, top), outward-pushing (Exp 2, middle), and large outward-pushing (Exp 3, bottom). The raincloud plots show the magnitude of the applied perturbations both as a distribution (the ‘cloud’) and with jittered raw data (the ‘rain’).

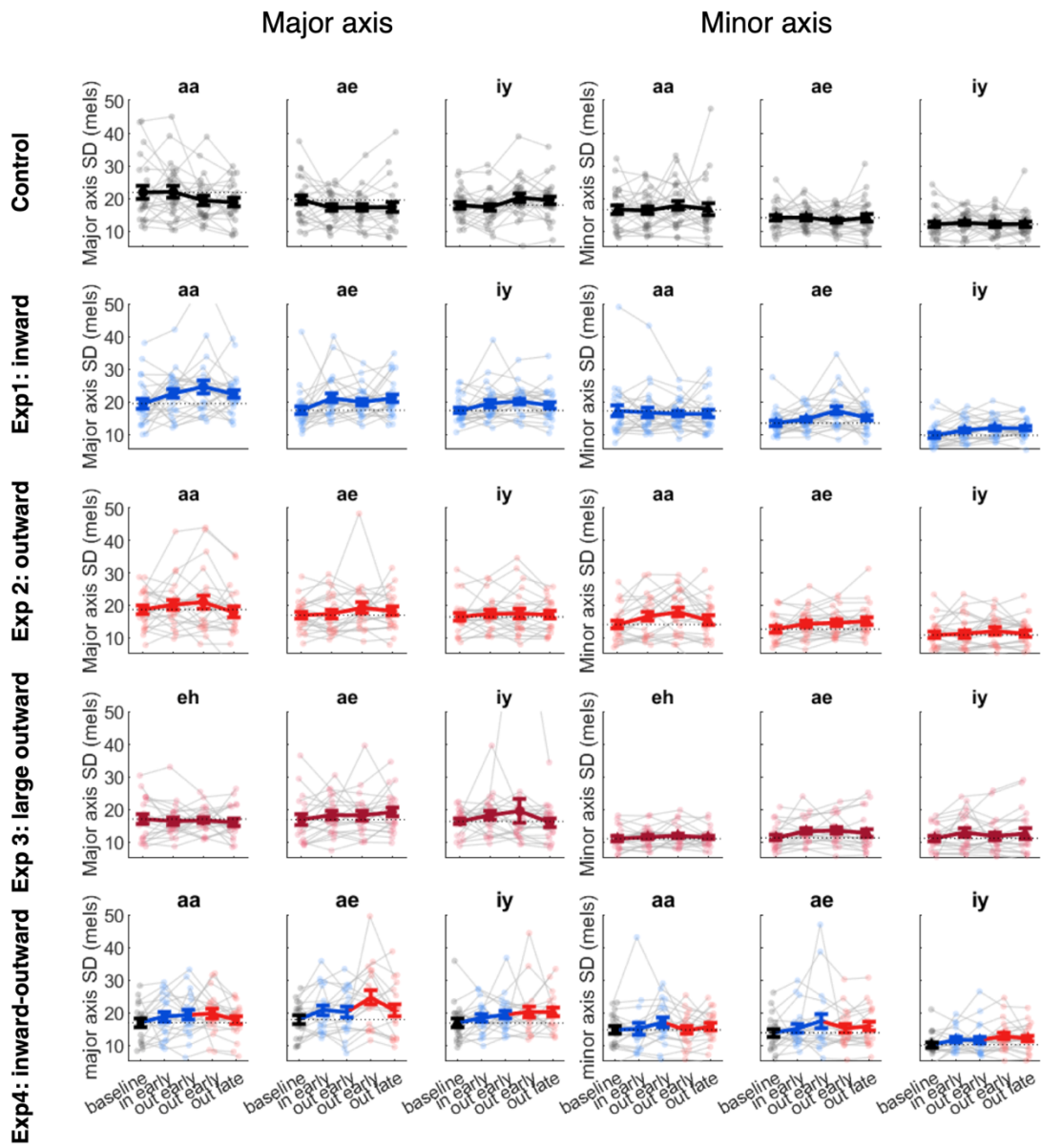
725
726
727
728
729
730
731
732
733
734
735

Table S1. Variability (SD) changes along major and minor axis

	Control	Experiment 1	Experiment 2	Experiment 3	Experiment 4
Main effect:					
Phase	F(2, 48) = 0.45, p = 0.642	F(2, 46) = 9.22, p < 0.001**	F(2, 40) = 5.74, p = 0.006*	F(1.5, 29.9) = 1.36, p = 0.267	F(2, 38) = 15.41, p < 0.001**
Vowel	F(2,48) = 5.39, p = 0.008*	F(2, 46) = 17.31, p < 0.001**	F(2, 40) = 17.82, p < 0.001**	F(1.3, 25.2) = 1.28, p = 0.28	F(2, 38) = 7.47, p = 0.002*
Axis	F(1, 24) = 110.13, p < 0.001**	F(1, 23) = 116.54, p < 0.001**	F(1, 20) = 40.02, p < 0.001**	F(1, 20) = 61.5, p < 0.001**	F(1, 19) = 77.23, p < 0.001**
Two-way interaction:					
Phase × Vowel	F(4, 96) = 1.90, p = 0.117	F(4, 92)=0.60, p = 0.662	F(4, 80)=2.01, p = 0.101	F(2.1, 41.4)=0.77, p = 0.474	F(4, 76)=1.06, p = 0.385
Phase × Axis	F(2, 48) = 0.83, p = 0.443	F(2, 46) = 2.84, p = 0.069	F(1.6, 31.5)=0.33, p = 0.669	F(1.5, 29.4)=0.28, p = 0.693	F(2, 38)=0.78, p = 0.465
Vowel × Axis	F(2, 48) = 5.87, p = 0.005*	F(2, 46) = 4.21, p = 0.021*	F(1.5, 29.2)=0.183, p = 0.185	F(1.2, 24)=0.05, p = 0.858	F(2, 38)=8.93, p < 0.001*
Three-way interaction:					
Phase × Vowel × Axis	F(4, 96)=1.69, p = 0.160	F(2.4, 54)=3.07, p = 0.047*	F(4, 80)=0.58, p = 0.681	F(2.2, 44)=2.1, p = 0.13	F(4, 76)=0.771, p = 0.547

Note. Significant three-way interaction was found only in Experiment 1. To follow up the significant three-way interaction, we grouped the data by vowel, and analysed the sample two-way interaction between phase and axis. There was no significant simple two-way interaction between phase and axis for any of the vowels (bod: p = 0.055, bad: p = 0.051, bead: p = 0.6). *p < 0.05 **p < 0.001.

736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754



755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772

Figure S2. Variability changes along major and minor axes in different words. Individual (small transparent dots, thin lines) and group means (large solid dots, thick lines) of formant variability in the baseline, exposure and washout phases. Error bars show standard error.

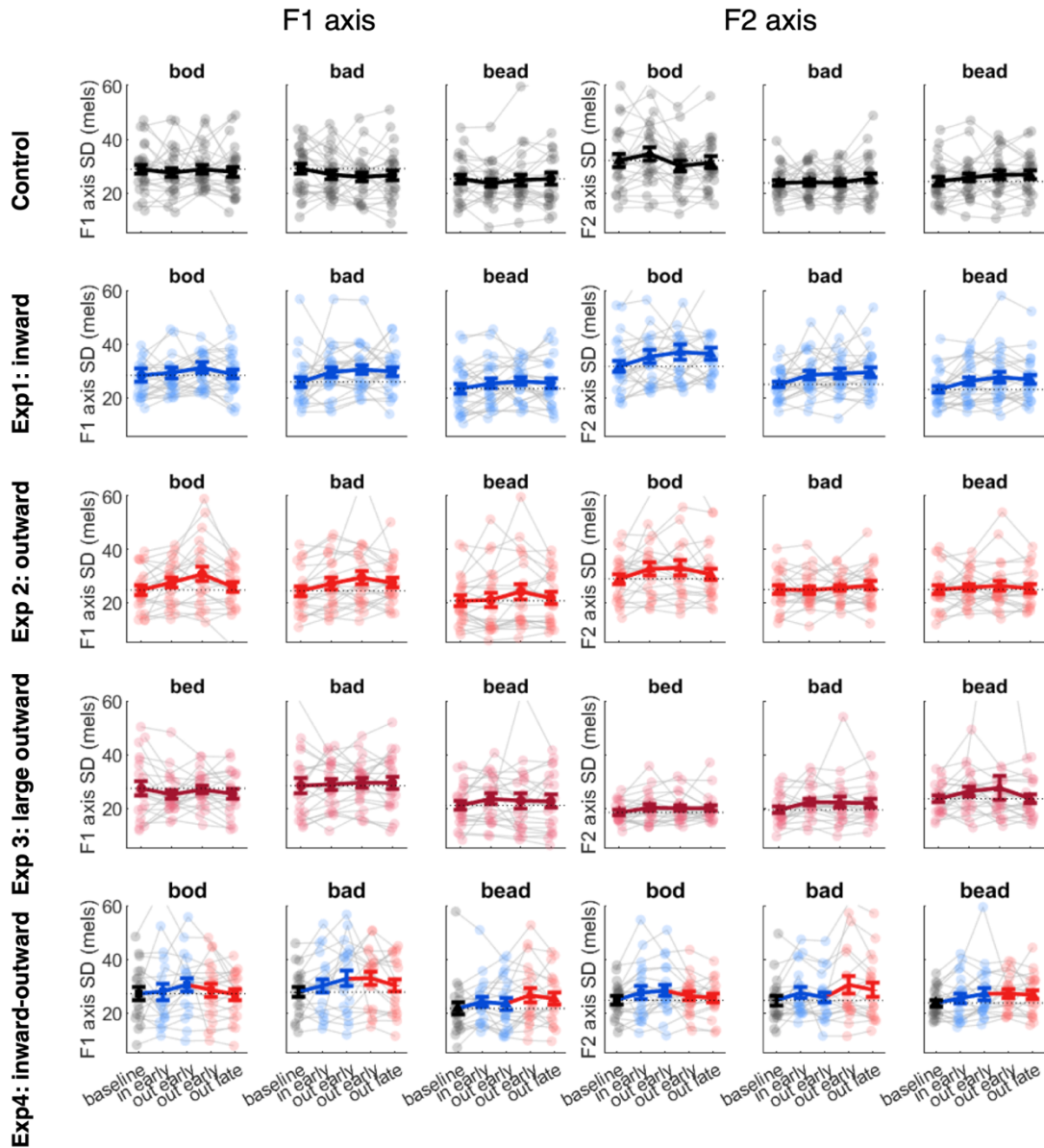
773
774
775
776
777
778

Table S2. Variability (SD) changes along F1 and F2 axis

	Control	Experiment 1	Experiment 2	Experiment 3	Experiment 4
Main effect:					
Phase	F(2, 48) = 0.25, p = 0.784	F(2, 46) = 7.38, p = 0.002*	F(2, 40) = 7.26, p = 0.002*	F(1.4, 28.1) = 1.30, p = 0.278	F(2, 38) = 12.02, p < 0.001**
Vowel	F(2,48) = 5.35, p = 0.008*	F(2, 46) = 15.30, p < 0.001**	F(2, 40) = 17.72, p < 0.001**	F(1.3, 27.1) = 1.68, p = 0.209	F(2, 38) = 5.66, p = 0.007*
Axis	F(1, 24) = 0.04, p = 0.837	F(1, 23) = 3.44, p = 0.077	F(1, 20) = 2.92, p = 0.103	F(1, 20) = 7.37, p < 0.013*	F(1, 19) = 1.43, p = 0.247
Two-way interaction:					
Phase × Vowel	F(4, 96) = 0.99, p = 0.417	F(4, 92) = 0.49, p = 0.744	F(4, 80) = 1.72, p = 0.153	F(2, 39.8) = 0.44, p = 0.644	F(4, 76) = 1.36, p = 0.256
Phase × Axis	F(1.6, 38) = 2.46, p = 0.110	F(2, 46) = 1.93, p = 0.161	F(2, 40) = 2.81, p = 0.072	F(2, 40) = 0.90, p = 0.416	F(2, 38) = 0.10, p = 0.907
Vowel × Axis	F(1.3, 32.8) = 2.40, p = 0.124	F(1.4, 31.5) = 3.30, p = 0.067	F(2, 40) = 3.35, p = 0.041*	F(1.5, 30.7) = 27.85, p < 0.001**	F(1.4, 27) = 1.73, p = 0.200
Three-way interaction:					
Phase × Vowel × Axis	F(4, 96) = 1.74, p = 0.148	F(2.4, 54) = 0.58, p = 0.678	F(4, 80) = 0.32, p = 0.862	F(4, 80) = 1.65, p = 0.17	F(4, 76) = 0.91, p = 0.463

Note. *p < 0.05 ** p < 0.001.

779
780
781
782
783
784



785
786
787
788
789
790
791

Figure S3. Variability changes along F1 and F2 axes in different words . Individual (small transparent dots, thin lines) and group means (large solid dots, thick lines) of formant variability in the baseline, exposure and washout phases. Error bars show standard error.

792 **References**

- 793 Abe, M. O., & Sternad, D. (2013). Directionality in distribution and temporal structure of
794 variability in skill acquisition. *Frontiers in Human Neuroscience*, 7.
795 <https://doi.org/10.3389/fnhum.2013.00225>
- 796 Ahn, J., Zhang, Z., & Sternad, D. (2016). Noise Induces Biased Estimation of the Correction
797 Gain. *PLOS ONE*, 11(7), e0158466. <https://doi.org/10.1371/journal.pone.0158466>
- 798 Baddeley, R. J., Ingram, H. A., & Miall, R. C. (2003). System Identification Applied to a
799 Visuomotor Task: Near-Optimal Human Performance in a Noisy Changing Task. *The*
800 *Journal of Neuroscience*, 23(7), 3066–3075. [https://doi.org/10.1523/JNEUROSCI.23-07-](https://doi.org/10.1523/JNEUROSCI.23-07-03066.2003)
801 [03066.2003](https://doi.org/10.1523/JNEUROSCI.23-07-03066.2003)
- 802 Blustein, D. H., Shehata, A. W., Kuylenstierna, E. S., Englehart, K. B., & Sensinger, J. W.
803 (2021). An analytical method reduces noise bias in motor adaptation analysis. *Scientific*
804 *Reports*, 11(1), 9245. <https://doi.org/10.1038/s41598-021-88688-5>
- 805 Boersma, P., & Weenink, D. (2019). *Praat: Doing phonetics by computer*. (Version 6.0.47)
806 [Computer software]. <http://www.praat.org/>
- 807 Burnett, T. A., Freedland, M. B., Larson, C. R., & Hain, T. C. (1998). Voice F0 responses to
808 manipulations in pitch feedback. *The Journal of the Acoustical Society of America*,
809 *103*(6), 3153–3161. <https://doi.org/10.1121/1.423073>
- 810 Cai, S., Boucek, M., Ghosh, S., Guenther, F. H., & Perkell, J. (2008). A System for Online
811 Dynamic Perturbation of Formant Trajectories and Results from Perturbations of the
812 Mandarin Triphthong /iau/. *Proceedings of the 8th International Seminar on Speech*
813 *Production*, 65–68.
- 814 Churchland, M. M., Afshar, A., & Shenoy, K. V. (2006). A Central Source of Movement
815 Variability. *Neuron*, 52(6), 1085–1096. <https://doi.org/10.1016/j.neuron.2006.10.034>

- 816 Daliri, A. (2021). A Computational Model for Estimating the Speech Motor System's Sensitivity
817 to Auditory Prediction Errors. *Journal of Speech, Language, and Hearing Research*,
818 64(6), 1841–1854. https://doi.org/10.1044/2021_JSLHR-20-00484
- 819 Daliri, A., & Dittman, J. (2019). Successful auditory motor adaptation requires task-relevant
820 auditory errors. *Journal of Neurophysiology*, 122(2), 552–562.
821 <https://doi.org/10.1152/jn.00662.2018>
- 822 Davids, K., Glazier, P., Araujo, D., & Bartlett, R. (2003). Movement Systems as Dynamical
823 Systems: The Functional Role of Variability and its Implications for Sports Medicine.
824 *Sports Medicine*, 33(4), 245–260. <https://doi.org/10.2165/00007256-200333040-00001>
- 825 Donchin, O., Francis, J. T., & Shadmehr, R. (2003). Quantifying Generalization from Trial-by-
826 Trial Behavior of Adaptive Systems that Learn with Basis Functions: Theory and
827 Experiments in Human Motor Control. *The Journal of Neuroscience*, 23(27), 9032–9045.
828 <https://doi.org/10.1523/JNEUROSCI.23-27-09032.2003>
- 829 Faisal, A. A., Selen, L. P. J., & Wolpert, D. M. (2008). Noise in the nervous system. *Nature*
830 *Reviews Neuroscience*, 9(4), 292–303. <https://doi.org/10.1038/nrn2258>
- 831 Harris, C. M., & Wolpert, D. M. (1998). Signal-dependent noise determines motor planning.
832 *Nature*, 394(6695), 780–784. <https://doi.org/10.1038/29528>
- 833 Herzfeld, D. J., & Shadmehr, R. (2014). Motor variability is not noise, but grist for the learning
834 mill. *Nature Neuroscience*, 17(2), 149–150. <https://doi.org/10.1038/nn.3633>
- 835 Houde, J. F., & Jordan, M. I. (1998). Sensorimotor Adaptation in Speech Production. *Science*,
836 279(5354), 1213–1216. <https://doi.org/10.1126/science.279.5354.1213>

- 837 Kang, N., Shinohara, M., Zatsiorsky, V. M., & Latash, M. L. (2004). Learning multi-finger
838 synergies: An uncontrolled manifold analysis. *Experimental Brain Research*, 157(3).
839 <https://doi.org/10.1007/s00221-004-1850-0>
- 840 Kewley-Port, D., & Watson, C. S. (1994). Formant-frequency discrimination for isolated English
841 vowels. *The Journal of the Acoustical Society of America*, 95(1), 485–496.
842 <https://doi.org/10.1121/1.410024>
- 843 Kim, K. S., & Max, L. (2020). Speech auditory-motor adaptation lacks an explicit component:
844 Reduced adaptation in adults who stutter reflects limitations in implicit sensorimotor
845 learning. *BioRxiv*, 2020.09.06.284638. <https://doi.org/10.1101/2020.09.06.284638>
- 846 Korzyukov, O., Bronder, A., Lee, Y., Patel, S., & Larson, C. R. (2017). Bioelectrical brain
847 effects of one’s own voice identification in pitch of voice auditory feedback.
848 *Neuropsychologia*, 101, 106–114.
849 <https://doi.org/10.1016/j.neuropsychologia.2017.04.035>
- 850 Nagamori, A., Laine, C. M., Loeb, G. E., & Valero-Cuevas, F. J. (2021). Force variability is
851 mostly not motor noise: Theoretical implications for motor control. *PLOS Computational*
852 *Biology*, 17(3), e1008707. <https://doi.org/10.1371/journal.pcbi.1008707>
- 853 Nault, D. R., & Munhall, K. G. (2020). Individual variability in auditory feedback processing:
854 Responses to real-time formant perturbations and their relation to perceptual acuity. *The*
855 *Journal of the Acoustical Society of America*, 148(6), 3709–3721.
856 <https://doi.org/10.1121/10.0002923>
- 857 Newell, K. M., & Vaillancourt, D. E. (2001). Dimensional change in motor learning. *Human*
858 *Movement Science*, 20(4–5), 695–715. [https://doi.org/10.1016/S0167-9457\(01\)00073-2](https://doi.org/10.1016/S0167-9457(01)00073-2)

- 859 Niziolek, C. A., & Guenther, F. H. (2013). Vowel Category Boundaries Enhance Cortical and
860 Behavioral Responses to Speech Feedback Alterations. *The Journal of Neuroscience*,
861 33(29), 12090–12098. <https://doi.org/10.1523/JNEUROSCI.1008-13.2013>
- 862 Niziolek, C. A., & Houde, J. (2015). *Wave_Viewer: First Release*.
863 <https://doi.org/10.5281/ZENODO.13839>
- 864 Niziolek, C. A., & Kiran, S. (2018). Assessing speech correction abilities with acoustic analyses:
865 Evidence of preserved online correction in persons with aphasia. *International Journal of*
866 *Speech-Language Pathology*, 20(6), 659–668.
867 <https://doi.org/10.1080/17549507.2018.1498920>
- 868 Niziolek, C. A., & Parrell, B. (2021). Responses to Auditory Feedback Manipulations in Speech
869 May Be Affected by Previous Exposure to Auditory Errors. *Journal of Speech,*
870 *Language, and Hearing Research*, 64(6S), 2169–2181.
871 https://doi.org/10.1044/2020_JSLHR-20-00263
- 872 Parrell, B., & Niziolek, C. A. (2021). Increased speech contrast induced by sensorimotor
873 adaptation to a nonuniform auditory perturbation. *Journal of Neurophysiology*, 125(2),
874 638–647. <https://doi.org/10.1152/jn.00466.2020>
- 875 Purcell, D. W., & Munhall, K. G. (2006). Compensation following real-time manipulation of
876 formants in isolated vowels. *Journal of Acoustical Society of America*, 119(4), 2288–
877 2297.
- 878 Renart, A., & Machens, C. K. (2014). Variability in neural activity and behavior. *Current*
879 *Opinion in Neurobiology*, 25, 211–220. <https://doi.org/10.1016/j.conb.2014.02.013>

- 880 Robertson, E. M., & Miall, R. C. (1997). Multi-joint limbs permit a flexible response to
881 unpredictable events. *Experimental Brain Research*, *117*(1), 148–152.
882 <https://doi.org/10.1007/s002210050208>
- 883 Scheerer, N. E., Behich, J., Liu, H., & Jones, J. A. (2013). ERP correlates of the magnitude of
884 pitch errors detected in the human voice. *Neuroscience*, *240*, 176–185.
885 <https://doi.org/10.1016/j.neuroscience.2013.02.054>
- 886 Scheidt, R. A., Dingwell, J. B., & Mussa-Ivaldi, F. A. (2001). Learning to Move Amid
887 Uncertainty. *Journal of Neurophysiology*, *86*(2), 971–985.
888 <https://doi.org/10.1152/jn.2001.86.2.971>
- 889 Scholz, J. P., & Schöner, G. (1999). The uncontrolled manifold concept: Identifying control
890 variables for a functional task. *Experimental Brain Research*, *126*(3), 289–306.
891 <https://doi.org/10.1007/s002210050738>
- 892 Shadmehr, R., Huang, H. J., & Ahmed, A. A. (2016). A Representation of Effort in Decision-
893 Making and Motor Control. *Current Biology*, *26*(14), 1929–1934.
894 <https://doi.org/10.1016/j.cub.2016.05.065>
- 895 Stein, R. B., Gossen, E. R., & Jones, K. E. (2005). Neuronal variability: Noise or part of the
896 signal? *Nature Reviews Neuroscience*, *6*(5), 389–397. <https://doi.org/10.1038/nrn1668>
- 897 Sternad, D., Abe, M. O., Hu, X., & Müller, H. (2011). Neuromotor Noise, Error Tolerance and
898 Velocity-Dependent Costs in Skilled Performance. *PLoS Computational Biology*, *7*(9),
899 e1002159. <https://doi.org/10.1371/journal.pcbi.1002159>
- 900 Thoroughman, K. A., & Shadmehr, R. (2000). Learning of action through adaptive combination
901 of motor primitives. *Nature*, *407*(6805), 742–747. <https://doi.org/10.1038/35037588>

- 902 Todorov, E. (2004). Optimality principles in sensorimotor control. *Nature Neuroscience*, 7(9),
903 907–915. <https://doi.org/10.1038/nn1309>
- 904 Tourville, J. A., Cai, S., & Guenther, F. (2013). Exploring auditory-motor interactions in normal
905 and disordered speech. *Proceedings of Meetings on Acoustics*, 19, 060180.
906 <https://doi.org/10.1121/1.4800684>
- 907 Tumer, E. C., & Brainard, M. S. (2007). Performance variability enables adaptive plasticity of
908 ‘crystallized’ adult birdsong. *Nature*, 450(7173), 1240–1244.
909 <https://doi.org/10.1038/nature06390>
- 910 van Beers, R. J., Brenner, E., & Smeets, J. B. J. (2013). Random walk of motor planning in task-
911 irrelevant dimensions. *Journal of Neurophysiology*, 109(4), 969–977.
912 <https://doi.org/10.1152/jn.00706.2012>
- 913 Wong, J., Wilson, E. T., Malfait, N., & Gribble, P. L. (2009). The Influence of Visual
914 Perturbations on the Neural Control of Limb Stiffness. *Journal of Neurophysiology*,
915 101(1), 246–257. <https://doi.org/10.1152/jn.90371.2008>
- 916 Wu, H. G., Miyamoto, Y. R., Gonzalez Castro, L. N., Ölveczky, B. P., & Smith, M. A. (2014).
917 Temporal structure of motor variability is dynamically regulated and predicts motor
918 learning ability. *Nat Neurosci*, 17(2), 312–321. <https://doi.org/10.1038/nn.3616>
919