Control of tongue movements by the Purkinje cells of the cerebellum

Paul Hage, Mohammad Amin Fakharian, Alden M. Shoup, Jay S. Pi, Ehsan Sedaghat-Nejad, Simon P. Orozco, In Kyu Jang, Vivian Looi, Hisham Y. Elseweifi, Nazanin Mohammadrezaei, Alexander N. Vasserman, Toren Arginteanu, and Reza Shadmehr

Laboratory for Computational Motor Control, Dept. of Biomedical Engineering, Johns Hopkins School of Medicine, Baltimore, Maryland 21205, USA

Correspondence: Reza Shadmehr, Johns Hopkins School of Medicine, 410 Traylor Building, 720 Rutland Ave., Baltimore, MD 21205. Email: shadmehr@jhu.edu

Acknowledgements: The work was supported by grants from the National Science Foundation (CNS-1714623), and the NIH (R01-EB028156, R37-NS128416).

Author contributions: PH and RS conceived experiments. PH, MAF, AMS, JSP, ESN, and SPO collected the data. PH, MAF, IKJ, VL, HYE, NM, ANV, TA, and RS analyzed the data. PH and HYE made the figures and performed statistical analysis. RS wrote the manuscript.

Competing interests: None.

Data availability: https://osf.io/WDXU4/

ORCID ID: Paul Hage 0000-0002-7006-9877, In Kyu Jang 0000-0002-1832-3743, Vivian Looi 0000-0002-9392-4178, Mohammad Amin Fakharian 0000-0002-7741-8689, Simon Orozco 0000-0001-6092-7228, Jay Pi 0000-0003-4036-2877, Ehsan Sedaghat-Nejad 0000-0001-6732-1296, Hisham Y. Elseweifi 0009-0005-8008-5829, Nazanin Mohammadrezaei 0009-0007-0549-9946, Alexander N. Vasserman 0009-0002-9794-1984, Toren Arginteanu 0000-0002-6029-7323, Reza Shadmehr 0000-0002-7686-2569.

Abstract

We use our tongue much like our hands: to interact with objects and transport them. For example, we use our hands to sense properties of objects and transport them in the nearby space, and we use our tongue to sense properties of food morsels and transport them through the oral cavity. But what does the cerebellum contribute to control of tongue movements? Here, we trained head-fixed marmosets to make skillful tongue movements to harvest food from small tubes that were placed at sharp angles to their mouth. We identified the lingual regions of the cerebellar vermis and then measured the contribution of each Purkinje cell (P-cell) to control of the tongue by relying on the brief but complete suppression that they experienced following an input from the inferior olive. When a P-cell was suppressed during protraction, the tongue's trajectory became hypermetric, and when the suppression took place during retraction, the tongue's return to the mouth was slowed. Both effects were amplified when two P-cells were simultaneously suppressed. Therefore, suppression of P-cells in the lingual vermis disrupted the forces that would normally decelerate the tongue as it approached the target. Notably, the population simple spike activity peaked near deceleration onset when the movement required precision (aiming for a tube), but not when the movement was for the purpose of grooming. Thus, the P-cells appeared to signal when to stop protrusion as the tongue approached its target.

Introduction

We use our tongue to shape the air and generate sounds in order to communicate, and we use our tongue to evaluate food morsels and transport them through the oral cavity in order to eat. These skillful acts involve coordination of over 100 muscles (1), producing movements that are fundamental to our existence. Damage to the cerebellum disrupts these movements, resulting in abnormal muscle activation patterns (2) that bear a resemblance to ataxia of the arm (3). However, life without a cerebellum in humans (4), or inactivation of the deep cerebellar nuclei in mice (5), do not eliminate tongue movements. Rather, the movements become inaccurate. For example, if the activities of Purkinje cells (P-cell) are disrupted via silencing of molecular layer interneurons, the tongue's trajectory becomes erratic and the subject is no longer able to efficiently harvest liquid rewards (6). Thus, the role of the cerebellum in control of the tongue may be similar to its function during control of the limbs (7) and the eyes (8, 9): stopping the movement on target. But how might the cerebellum achieve this?

In primates, stimulation of the fastigial nucleus moves the tongue predominantly in the ventral-dorsal axis, while stimulation of the dentate nucleus moves it mainly in the medial-lateral axis (10). Notably, tongue muscles are most readily activated via stimulation of the fastigial nucleus (as compared to the other cerebellar nuclei) (11), suggesting that the P-cells in the vermis play a prominent role in control of the tongue. Unfortunately, there are no reports of P-cell activity in the vermis during targeted tongue movements in any species, but more is known regarding activity of P-cells in Crus I and Crus II (in rodents). For example, as a licking bout is about to start, many P-cells in Crus I and Crus II increase their simple spikes (SS), while a smaller number exhibit a decrease (12). Once the licking begins, the SS rates as a population are phase-locked to the rhythm of the lick, with peaks occurring near lick onset (5). Complex spikes (CSs) also exhibit their highest rates during protraction (12, 13). However, it has been difficult to understand the relationship between the activities of P-cells and control of the tongue.

To answer this question, we sought an animal model that had a long tongue and could skillfully direct it to small targets. Marmosets are an attractive choice because they have a 21mm tongue which they use to burrow into small holes and retrieve insects and sap (14). Indeed, they have an extraordinary ability to control their tongue, vocalizing in order to label other marmosets during 2-way communication (15).

As we trained head-fixed marmosets to make saccades to visual targets and then rewarded them with food (16), we noticed that they could naturally bend and twist their tongue in order to insert it into small tubes, even when the tubes were placed at 90° with respect to their mouth (17). Because their harvest was difficult, they chose to do many saccade trials, allowing the food to accumulate, then stopped working and claimed their cache by scooping the food out of the tube (17).

To quantify how the cerebellum was contributing to the control of the tongue, we recorded from tongue modulated P-cells in lobule VI and VII of the vermis. Then, we relied on the fact that the inferior olive not only transmitted unexpected sensory events to the cerebellum (8, 18–20), it also acted as a stochastic perturbation that completely suppressed the P-cells (21, 22), which then resulted in a small movement (23), or a disruption of the ongoing movement (24–26). Using spike-triggered averaging on the climbing fiber input, we found that the resulting SS suppression altered the deceleration phase as the tongue approached the target, producing hypermetria.

This hypermetria was replicated when the P-cells experienced a long period of SS pause without a preceding CS. That is, both a CS-induced SS suppression, and a long SS pause independently had the

same effect on behavior: producing downstream forces that extended the tongue. Because as a population, the SS rates were greatest during the deceleration phase of protraction, the results suggested that the P-cells signaled downstream structures to stop the movement as the tongue approached its target. Indeed, this strong engagement of the P-cells was present when the tongue was aiming for a small tube, but not when the movement's purpose was to groom the face.

Results

We trained marmosets (n=3) to perform visually guided saccades in exchange for food (Fig. 1A). The subjects performed a sequence of task-relevant saccades, at the end of which we delivered an increment of food (slurry mixture of apple sauce and monkey chow). This food was presented via either the left or the right tube for 50-300 consecutive trials, and then switched tubes. Because the food amounts were small (0.015-0.02 mL), and the tubes were located at ±90° with respect to the mouth, the harvest was effortful (17), requiring skillful movements toward a target that was just large enough to accommodate the tongue (4.4 mm diameter tube). As a result, the subjects chose to work for a few consecutive saccade trials (n=6.1±0.02 successful trials per work period), allowing the food to accumulate, then stopped making saccades to targets, fixated the tube and harvest via a bout of licking (n=22.03 ± 0.04 licks per harvest period, Fig. 1B).

We tracked the motion of the tongue in the horizontal plane using DeepLabCut (27). The tongue movements were of two general types: in the task relevant licks the subjects aimed for the tube (**video 1**), whereas in the task-irrelevant licks the subjects groomed their mouth (Fig. 1A) (**video 2**, **video 3**, **video 4**). During the 2-3 hour recording sessions the subjects performed n=4401±11 task relevant licks (i.e., aimed at a tube, mean±SEM), and n=1310±4 task irrelevant licks (Fig. 1E).

Typically, the subjects began their harvest by licking the food near the tip of the tube (**video 5**), but then as the food cache declined, they inserted their tongue into the tube (**video 6**, **video 7**), scooping out their reward. Thus, we divided the task relevant licks into two subtypes, those that aimed for the edge of the tube and harvested the food that was near the tip (Fig. 1A, labeled 2&4), and those that penetrated the tube and harvested the food that was deeper (Fig. 1A, labeled 1&5). Lick protraction velocity was largest for inner tube licks, which also had the largest amplitude and longest protraction duration (Supplementary Fig. S1). Duration of the protraction phase of the inner tube licks was longer than the duration of retraction. For example, in subject 132F, inner tube licks had a protraction duration of 201.3+/-0.16 ms vs. retraction duration of 133.6 +/-0.12 ms (Supplementary Fig. S1). This is consistent with the idea that in contrast to retraction, protraction required aiming which tended to accompany longer duration movements.

Climbing fibers were most active near lick onset

We combined MRI and CT image guided procedures (16) to place heptodes and silicon probes in lobules VI and VII of the vermis. Over the course of 3.5 years, we recorded from n=284 P-cells (Figs. 1C & 1E) while the subjects performed 840,787 licks. A neuron was identified as a definitive P-cell (n=230) because of the presence of complex spikes (CS). In addition, we included data from putative P-cells (n=54) for which we could not isolate the CSs but the neuron was located in the P-cell layer and exhibited 0 ms synchronous simple spike (SS) interactions with other confirmed P-cells (8, 9, 26, 28) (Supplementary Fig. S2A).

Among our P-cells, the SS modulations were usually present for both tongue movements and eye movements (Supplementary Fig. S2C). However, as our aim was to record from the P-cells that were tongue modulated, in our population the P-cells were more strongly modulated by tongue movements (paired t-test, t(156)=7.96, p=3.3E-13). This preferential encoding of tongue vs. eye was greater for neurons that were located in lobule VI (Supplementary Fig. S2B).

Fig. 1D illustrates the activities of two P-cells near bout onset, as well as during licking. As the bout began, one P-cell increased its SS activity, earlier when the tongue targeted the ipsilateral tube (with respect to the site of recording), while another P-cell decreased its SS activity, earlier for the contralateral tube. As the licking continued, the SS rates in both P-cells were modulated in a rhythmic pattern. We focused on three periods during each lick: protraction acceleration period (Fig. 1D, a-b), protraction deceleration period (b-c), and retraction acceleration period (c-d).

As a population, the n=230 confirmed P-cells exhibited CS rates that increased near the onset of protraction (Figs. 2A), peaking around the time the tongue touched the tube, and then decreased below baseline around retraction onset. The increased rates during protraction were larger for ipsilateral licks (within cell difference, protraction period, mean±SEM: 0.077±0.024, t(229)=26.7, p=1.9E-72). Thus, as a population, for both target directions, the phase of movement for which the CS rates were maximum was protraction (termed CS-on phase).

P-cell suppression produced overshooting during protraction and slowed return during retraction

The climbing fiber input suppressed SS production (Fig. 2B, left: all P-cells, right: single P-cell), lasting an average of 14.8±0.44 ms during licking (time to 85% recovery of SS rate, mean±SEM). However, the CS events were rare: a CS occurred in only 5.45±0.01% of the licks during protraction acceleration period (mean±SEM), 9.28±0.01% of the licks during protraction deceleration period, and 6.59±0.009% of the licks during retraction acceleration period (Fig. 2C). We collected a large number of licks per neuron (4401±11 licks, Fig. 1E), then performed CS-triggered averaging to ask whether the resulting SS suppression affected the motion of the tongue.

For each P-cell we considered triplets of consecutive licks $\{n-1, n, n+1\}$ in which all three licks were of the same type, i.e., contacted the same part of the tube (edge or inner). We then selected those triplets in which there was a CS at only a single period in lick n, but no CS during any period in the two neighboring licks $n-1$, and $n+1$. For example, consider licks in which there was a CS in one P-cell during the acceleration period of protrusion (Supplementary Fig. S3). This acceleration period was brief (49.9±0.019 ms), during which the SS activity was normally rising and nearly identical in licks $n-1$ and $n+1$ but suppressed during lick n (Supplementary Fig. S3, 1st row). We measured the change in SS activity by comparing lick n with licks $n-1$ and $n+1$ and plotted the results of each comparison in Fig. S3B, green & blue solid lines (labeled suppressed, top row, right column). As a control, we performed a bootstrapping procedure in which we generated a pseudo data set for each P-cell by randomly assigning the CS label to a lick and comparing it with its two temporally adjacent neighbors.

To ask whether there were any effects of the brief SS suppression on the tongue, we measured the distance of the tongue tip to the mouth and also its angle with respect to the midline, and then compared the trajectories in lick n with the neighboring licks in which the P-cell was not suppressed. The effects appeared consistent (Supplementary Fig. S3A, red colors): following SS suppression, there was little or no change in tongue kinematics. Statistical testing, which relied on bootstrapping procedure to

compute 95% confidence intervals (CI), demonstrated that the changes were within the error bounds. Thus, the suppression during the acceleration period of protraction had no significant effects on tongue trajectory (Supplementary Fig. S3C, measured via distance to the mouth, and its angle, at peak protraction speed and peak displacement).

As the protraction continued, the effects of SS suppression became evident. If the suppression occurred during the deceleration period of protraction (duration: 103.6±0.04 ms), the tongue exhibited hypermetria (Fig. 3A, second row), producing increased displacement and increased angle of the tongue's trajectory (Figs. 3B, displacement: 0.37±0.002 mm, 95%CI = [-0.17 0.14] mm, angle: 3.35±0.02 deg, 95%CI = [-1.14 0.98] deg). Notably, the effects were consistent regardless of whether lick n was compared to the previous lick ($n-1$), or the subsequent lick ($n+1$) (Fig. 3B, second row, blue and green traces). Furthermore, the effects were consistent across the P-cells (Fig. 3A, red colors, also Supplementary Fig. S4, left panel), and larger for contralateral licks (within cell difference, ipsilateral minus contralateral, at lick endpoint, displacement mean±SEM: -0.15±0.04 mm, t(229)=-3.97, p=9.6E-05, angle: -1.41±0.37, t(229)=-3.79, p=9.69E-05).

These results hinted that SS suppression prevented normal deceleration, producing hypermetria and a bending of the tongue away from the midline. If this interpretation is valid, then a similar suppression during the retraction period should produce forces that are again in the direction of protraction, now resisting the tongue's return. That is, if SS suppression during protraction sped the movement outward, then the same suppression during retraction should now slow the movement. In both cases, the suppression should bend the tongue away from the midline.

As retraction began, the population CS activity (Fig. 2A) had fallen below baseline, i.e., opposite of the activity during protraction. Yet, if the CS occurred during the retraction acceleration period (duration: 78.3±0.02 ms), the resulting suppression was again an outward displacement of the tongue and bending (Fig. 3D). A comparison of lick n with $n-1$ or $n+1$ revealed a consistent effect: an increased distance of the tongue to the mouth and an increased angle (Fig. 3F, displacement: 1.29±0.002 mm, 95%CI = [-0.32 0.22] mm, angle: 4.70±0.02 deg, 95%CI = [-1.15 0.85] deg). These effects were present for ipsilateral and contralateral licks (Fig. 3E), larger for the contralateral licks (within cell difference, ipsilateral minus contralateral, at lick endpoint, displacement: -0.37±0.09 mm, t(229)=-3.97, p=4.8E-05, angle: -2.06±0.50, t(229)=-4.1, p=2.71E-05), and consistent across the P-cells (Fig. 3D, red colors). Thus, the CS-induced SS suppression during retraction slowed the return of the tongue and producing bending away from the midline.

In summary, when the climbing fiber input briefly suppressed the P-cells during the deceleration period of the tongue's protrusion, the tongue overextended and bent away from the midline. When this suppression occurred during retraction, the tongue's return was slowed and again bent away from the midline. These effects were present for targets on both the ipsilateral and contralateral sides, but greater when the target was contralateral. Thus, it appeared that SS suppression disrupted the motor commands that would normally stop the tongue during protraction and return it during retraction. That is, the downstream effect of suppression of P-cells was to produce forces that extended the tongue and produced lateral bending.

Hypermetria increased when pairs of P-cells were simultaneously suppressed

Our dataset included n=298 pairs of simultaneously recorded P-cells. This allowed us to test whether near simultaneous suppression of two P-cells had a greater effect on tongue kinematics as compared to when only one of the two P-cells was suppressed.

As before, we collected triplets of consecutive licks $\{n-1, n, n+1\}$ where all the licks were of the same type (directed toward the same part of the tube), but only lick n had a CS. We divided the triplets based on whether a CS was present in only one of the P-cells, or both P-cells, then computed trajectory differences between licks n and $n-1$, as well as licks n and $n+1$. Finally, for each pair of P-cells we averaged $n-(n-1)$ and $n-(n+1)$ to increase statistical power (as there were far fewer licks in which both P-cells experienced a CS during the same period of the movement).

We found that if two P-cells were suppressed during the deceleration phase of protraction, then there was significantly greater displacement of the tongue, and bending, as compared to when only one of the P-cells was suppressed (Fig. 4A, angle: t(248) = -3.7167, p = 2.5E-04, displacement: t(248) = -5.5471,p = 7.43E-08). Similarly, when the suppression occurred during retraction, the return phase of the movement experienced a greater slowing in the case of two P-cells as compared to a single P-cell (Fig. 4B, angle: t(192) = 6.34, p = 1.6tE-09, displacement: t(192) = 6.50, p = 6.79E-10). Thus, near simultaneous suppression of two P-cells roughly doubled the kinematic effects.

Control studies

The fact that a CS was present during a given period in lick n may have been because earlier in the tongue's trajectory there was an event (for example, an error), that affected that movement, increasing the likelihood of a CS, and resulting in compensatory movements that followed the CS. To check for this, for each period during which we observed a CS we considered the tongue's trajectory in the same lick but during the period preceding the CS. For example, for the licks in which there was a CS in the protraction deceleration period we focused on the acceleration period of the same movement. By comparing the lick in which the CS had occurred with its two neighbors, we found that in the period before the CS had occurred tongue kinematics remained within the 95% confidence intervals of chance: distance to mouth and angle of the tongue at peak speed were not different than chance (Fig. 3C, period before peak speed, displacement: -0.11±0.001 mm, 95%CI = [-0.12 0.09] mm, angle: 0.19±0.01 deg, 95%CI = [-0.73 0.53] deg).

Next, we checked the licks in which there was a CS in the retraction period. We found that before the CS had occurred, at the onset of retraction the distance to the mouth and tongue angle were not different than if the CS had not occurred (Fig. 3F, protraction period before peak displacement, displacement: 0.01±0.002 mm, 95%CI = [-0.16 0.10] mm, angle: 0.68±0.02 deg, 95%CI = [-1.14 0.83] deg). Moreover, during the preceding protraction in the same lick, at peak tongue speed, the distance and angle were again not different than chance (displacement: -0.66±0.003 mm, 95%CI = [-0.70 0.74] mm, angle:-2.16±0.02 deg, 95%CI = [-3.35 3.58] deg). Yet, if during the return phase a CS was present, the movement was slowed (Fig. 3F, before peak-speed period of retraction, displacement: 1.29±0.002 mm, 95%CI = [-0.32 0.22] mm, angle: 4.70±0.02 deg, 95%CI = [-1.16 0.85] deg). That is, the tongue's trajectory before the CS was not significantly different than the neighboring licks, but after the CS-triggered SS suppression the trajectories diverged.

As a further control, we considered movements in which there was a CS event just before protrusion onset. In these movements, the SS rates were suppressed, but the rates recovered around 10

ms after lick onset (Supplementary Fig. S5, 1st row). Thus, despite the presence of a CS just before the lick had started, the SS rates during the entire protrusion and retraction periods were intact. The trajectory of the tongue as measured via distance and angle remained within the 95% confidence interval bounds of chance (Supplementary Fig. S5, 3rd and 4th rows).

Effect of SS suppression remained consistent across P-cells regardless of SS modulation

For nearly every P-cell in our dataset, during both protraction and retraction, the CS-triggered SS suppression was followed by an extension of the tongue and a lateral bending (Figs. 3A & 3D). This consistency was surprising in light of the diversity that was present in the SS patterns: before the onset of the bout, some P-cells had increased their SS rates with respect to baseline, while others had decreased (Fig. 1D). Did the effects of SS suppression differ in these two groups?

In our data set of n=142 of P-cells with SSs, most cells (n=123) increased their SS rates before the onset of the bout, but a minority exhibited a decrease (n=31) (Figs. 5A & 5B). We separated the P-cells into SS increasers and decreasers and found that despite the differences in their SS patterns, the CS rates increased in both groups near bout onset (Fig. 5C, peak CS firing rate change from baseline, SS increasers: 0.15±0.04 Hz, SS decreasers: 0.26±0.10 Hz). We then compared tongue trajectories in triplets of consecutive licks $\{n-1, n, n+1\}$, finding that during protraction, following SS suppression in the deceleration period, in both groups of P-cells the tongue was displaced away from the mouth, exhibiting a greater distance and angle (Fig. 5D, SS increasers, displacement: 0.38±0.005 mm, 95%CI = [-0.02 0.02] mm, angle: 3.72±0.05 deg, 95%CI = [-0.16 0.15] deg, SS decreasers: displacement: 0.45±0.04 mm, 95%CI = [0.04 -0.06] mm, angle: 4.28±0.54 deg, 95%CI = [-0.16 0.15] deg). Similarly, during retraction, SS suppression in both groups of P-cells produced a slowing of the tongue's return, resulting in a greater distance and angle (Fig. 5E, SS increasers, displacement: 1.25±0.01 mm, 95%CI = [-0.04 0.03] mm, angle: 4.45±0.05 deg, 95%CI = [-0.16 0.17] deg, SS decreasers: displacement: 1.04±0.04 mm, 95%CI = [-0.04 0.03] mm, angle: 3.48±0.17 deg, 95%CI = [-0.16 0.17] deg).

Thus, regardless of the patterns of SS activity during licking among the various P-cells, the downstream effect of CS-triggered SS suppression was extension of the tongue coupled with lateral bending.

SS pause without a CS was sufficient to produce hypermetria

We had interpreted the kinematic effects that followed the CS-induced SS suppression as being caused by SS suppression, not due to the presence of a CS. To check the validity of this conjecture, we quantified the kinematic effects of non-CS-induced long SS pauses on the tongue's trajectory. To identify a long pause, for each P-cell we considered all licks towards the same direction in which no CS occurred at any point in the movement. We then identified the longest inter-spike interval (ISI) for the SSs in each phase of each lick (phase refers to protraction deceleration period, etc.). For each phase under study, and each P-cell, we labeled 25% of the licks with the largest ISIs as a "long-pause" lick.

Next, we considered triplets of consecutive licks $\{n-1, n, n+1\}$ of the same type in the same direction in which none of the licks had a CS during any phase of the movement. We selected the subset of triplets in which lick n had a long pause in only one phase of the movement, but no long pauses in any phase of the neighboring licks. We then compared tongue trajectories between the lick that had a long SS pause with its two neighboring licks.

On average, the duration of a long SS pause was 31.25±1.2 ms during protraction deceleration, and 34.25±1.08 ms (mean±SEM) during retraction acceleration. If this pause occurred during the protraction acceleration period, it produced hypermetria and bending of the tongue away from the midline (Fig. 6A), and if it occurred during the retraction period it produced a slowing of the return and also a bending away from the midline (Fig. 6B). The trajectory changes were quite similar to what we had observed following a CS-induced SS suppression.

Thus, regardless of whether an SS pause was due to the arrival of a CS or not, the downstream effects were the same: production of forces that extended the tongue, bending it away from the midline.

The SS rates peaked at deceleration onset, but only if the movement was reward relevant

Across the cells, the CS rates peaked at approximately the time of maximum protraction velocity (Figs. 7A, left column), exhibiting a greater rate for movements toward the ipsilateral side (ipsilateral: 0.21±0.02 Hz, contralateral: 0.16±0.02 Hz, within cell difference, average CS rate, ipsilateral minus contralateral, t(229)=28.1, p=2.5E-76). In contrast, the CS rates declined below baseline before the onset of retraction. Thus, the CS-on phase across the P-cells was protraction.

To analyze the SS activities as a population, we had to consider the fact that during a bout, as the SS rates modulated about a mean, the mean was not stationary (Fig. 5A). Rather, the mean SS rates rose or fell before bout onset, then drifted back toward the values before start of the bout, reaching pre-bout levels at the bout ended. Thus, to quantify modulation of SS rates, for each P-cell we considered a 2 sec moving window to compute the running average of its firing rate, then computed the SS rates with respect to this mean. The 2 sec window was chosen because it was roughly an order of magnitude larger than the duration of a typical lick.

Given that a CS-triggered suppression in SS rates induced downstream forces that pushed the tongue outwards, then if this region of the cerebellum was interested in stopping the ongoing movement, during deceleration of protraction the SS rates should increase, thus commanding forces that would stop the motion of the tongue as it neared the target. Indeed, the SS rates peaked near the onset of lick deceleration, and were larger for contralateral licks (Fig. 7A, right column, ipsilateral: 11.58±1.75 Hz, contralateral: 12.91±1.95 Hz, within cell difference: t(156)=13.6, p=2.9E-28).

To test if SS modulation varied with tongue kinematics, we considered two conditions: when licks had the same amplitude but different peak velocities (Fig. 7B), and when they had the same peak velocity but different amplitudes (Fig. 7C). To consider licks of the same amplitude but different velocity, we quantified lick vigor, defined as the peak speed of the protraction with respect to the speed expected for a lick of the same amplitude (17, 29) (Supplementary Fig. S6). High vigor licks exhibited greater SS rates at peak protraction speed (Fig. 7B, within cell difference, high vigor minus low vigor, 2.92±0.60 Hz, t(156)=14.1, p=1.2E-29). That is, licks that required greater deceleration forces accompanied greater SS rates near the onset of deceleration.

Next, we considered licks that had the same peak velocity but different amplitudes (Fig. 7C). These two licks began with very similar velocity patterns, reaching on average identical peak velocity, but the SS rates achieved a greater peak rate for the licks that had a longer deceleration period. Moreover, the licks with the longer period of deceleration had a longer period of SS rate increase (Fig. 7C, time to baseline crossing, high duration: 126.80±7.68 ms, low duration: 86.02±7.23 ms, paired t-test,

t(156)=7.84, p=6.75E-13). As a result, the SS firing rates at peak velocity tended to increase with the lick's peak amplitude (Fig. 7E).

All of these results were for tongue movements in which the subject aimed for the small food tubes. Would the same patterns hold for movements that did not require such precision? To consider this question, we turned to the licks that were not directed toward a food tube, which were often grooming licks. These licks tended to be slower, with a peak speed that was roughly half the speed of the licks directed toward the food tubes (Fig. 7D, peak velocity of food tube licks: 290.5±3.6 mm/s, other licks: 153.8±2.5 mm/s). Remarkably, during protraction of these licks the CS rates were an order of magnitude smaller than when the licks were toward the food tubes (Fig. 7D, paired t-test, tube licks vs. grooming, combined directions, t(229)=-3.27, p=0.00126). Similarly, the SS rate modulations were much smaller (Fig. 7D, right, paired t-test, tube vs. other, combined directions, t(156)=8.6, p=8.44E-15). In sharp contrast, during retraction of these task-irrelevant licks, both the CS and the SS rates were modulated below baseline by amounts that were roughly comparable to the rates of the tube directed licks. That is, the fundamental difference in the P-cell activity between the tube directed and other licks was in the protraction phase, the phase in which control of the tongue required precision.

In summary, when the licks were directed toward the food tube, the CS and SS rates peaked during protraction and were greater when the movement had greater speed. Because SS suppression produced forces that extended the tongue, the fact that SS activity peaked near deceleration onset was consistent with the view that the downstream effects were to decelerate the tongue as it neared its target. Remarkably, both the CS and SS modulation patterns were largely absent when the tongue movements were not aimed at the food tube.

Error in the tongue's trajectory was reported to the cerebellum via complex spikes

Inserting the tongue into the tube required precision because the opening was only slightly larger than the tongue. As a result, in roughly 15% of the licks the food was inside the tube, but the tongue missed the entrance (**video 8**, **video 9**, and **video 10**, also Supplementary Fig. S7). We labeled these as unsuccessful licks because the tongue did not bend enough and instead hit the tube's outer edge. Were these errors reported to the cerebellum?

To visualize the CS patterns as a function of the spatial location of the tongue, we needed to compute $\Pr(CS|x)$, i.e., the probability of producing a complex spike given that the tongue was at a given location. To arrive at this variable, we began with computing $p(x|CS)$, i.e., the probability density of the position of the tongue's tip x, given that a CS occurred at time t. To compute this function, we used spike triggered averaging to compute the average position of the tongue during the 50ms period before the CS. We found that this likelihood depended on whether the tongue successfully entered the tube or not. For licks that were successful and entered the tube, the likelihood $p(x_s|CS)$ was bimodal, exhibiting a peak near lick onset, then a second peak near the food (Fig. 8A, left). For licks that were unsuccessful and did not enter the tube, the likelihood $p(x_u|CS)$ was also bimodal, but now the second peak was around the edge where the tongue had collided with the tube (Fig. 8B, left).

We next computed the marginal probability density $p(x)$ for the successful and unsuccessful licks (Fig. 8A and 8B, right). This function estimated the probability of the tongue being at a given position during the various licks. We then computed the prior probabilities $\Pr(CS)$, and the ratio of the probabilities $p(x|CS)\Pr(CS)/p(x)$, arriving at the posterior probability $\Pr(CS|x_s)$ and $\Pr(CS|x_u)$. Each

posterior estimated the probability of observing a CS as a function of the location of the tongue during successful and unsuccessful licks. Finally, we computed the error-induced spatial pattern of complex spikes by subtracting the posteriors (Fig. 8C). The results revealed a spatial gradient that increased with the tongue's distance from the mouth, suggesting that once we accounted for the CS modulations associated with normal licking, the CS events that remained in the unsuccessful licks tended to occur after the tongue had touched the far end of the tube.

To view these error-specific effects in another way, we plotted the CS rates as a function of time with respect to the tube-touch event. For the successful licks, tube touch occurred when the tongue crossed the tube's edge and entered the tube. In this case, the CS rates following tube touch were depressed (Fig. 8D). In contrast, for the unsuccessful licks the tube touch indicated an error, and the CS rates following this event showed a sharp increase (100 ms period following tube touch, CS rate in unsuccessful vs. successful licks, paired t-test, $t(459)=8.09$ $p=5.3\text{x}10\text{E-}15$).

Thus, when the food was inside the tube and the tongue successfully entered it, tube touch did not elicit complex spikes. However, when the food was inside the tube and the tongue touched it, but failed to enter it, now the tube touch event produced complex spikes.

Discussion

To quantify how the P-cells in the vermis contributed to control of the tongue, we trained marmosets to make skillful movements, extending and bending their tongue to harvest food from small tubes that were placed at 90° with respect to their mouth. Using spike-triggered averaging on the climbing fiber input to each P-cell, we found that if the resulting SS suppression occurred during protraction, there was a disruption in the deceleration phase of the movement, resulting in hypermetria as the tongue approached the target. A suppression that occurred during retraction retarded the tongue's return. When two P-cells were simultaneously suppressed, the kinematic effects magnified. Thus, regardless of whether a P-cell was suppressed during protraction or retraction, the downstream effects were production of forces that pulled the tongue outwards. These results were present for both ipsilateral and contralateral targets, but greater when the target was contralateral.

Because during unperturbed movements the SS rates in the population peaked as the movement began decelerating, we infer that the downstream effects were production of forces in the direction of retraction. This suggests that the contributions of the cerebellum to control of tongue movements may be similar to that of the eyes (30, 31) and the limbs (7): steering the movement and stopping it as it nears the target.

P-cells were modulated only for reward relevant movements

When the purpose of the movement was to groom the face, during the protraction phase the CS rates remained at baseline while the SS modulations were absent. This observation replicated our observations during saccades: when saccades are aimed toward a reward relevant target, in lobule VII the CS and SS rates are strongly modulated and the P-cell population predicts when the movement should be stopped (26). When similar saccades are made without a reward relevant target, the CS modulations are missing (8) and the SS rates no longer predict deceleration onset (26). Thus, for both saccades and licking, the cerebellum is engaged only when the movement is reward relevant.

One reason for engagement of the P-cells during reward relevant movements may be the greater accuracy requirements of those movements. Target position (for saccades) strongly engages the

neurons in the superior colliculus, which appear to transmit that information to the cerebellum via mossy fibers (26). When the saccade is reward irrelevant, the encoding of the target location is muted in the colliculus (32–34), as well as the mossy fibers (26). This implies that when that movement is not reward relevant, the cerebellum is poorly informed of the goal of the movement. As a result, it cannot assist in predicting when the movement is nearing the target and should be stopped.

Because the colliculus contains a topographic of map of tongue movements (35), we conjecture that like saccades, tongue movements that are not reward relevant will produce muted activity in the colliculus.

Climbing fibers were most active near movement onset

We were surprised that the CS rates peaked not after the tongue contacted the tube, but around the onset of protraction. This observation reproduced findings of Welsh et al. (13) in rats, who recorded from the lateral cerebellum and found that the CS rates peaked near the onset of tongue protrusion, even when the tongue was deafferented. Indeed, in many types of movements, including walking (36), reaching (37–41), and moving the wrist (42), the CS rates peak near movement onset. For example, in the oculomotor vermis, near saccade onset the olivary input informs the P-cells regarding the direction of the upcoming movement (8). How might the inferior olive be involved in transmitting movement information to the cerebellum?

A key observation is that the inferior olive not only receives input from the superior colliculus (43–45), but that subthreshold stimulation of a region of the colliculus leads to CS production without producing a movement (46). Thus, it is possible that the increased CS rates around the onset of protraction reflect activities of regions that initiated the tongue movement, i.e., the motor cortex and the superior colliculus (47). This prediction remains to be tested with simultaneous recordings from the colliculus, motor cortex, and the cerebellum.

Climbing fibers reported lick errors

Because the tubes were placed at sharp angles to the mouth, roughly 15% of the licks failed to retrieve the food. In these unsuccessful licks, the tongue did not bend enough and instead collided with the far edge of the tube. The climbing fibers reported this error robustly, exhibiting a strong increase in rates following the touch of the tube. Remarkably, when the touch event was not in error, i.e., the tongue entered the tube, the CS rates were suppressed. Thus, the CS rates signaled a touch that was in error, not a touch that was expected.

Using the olivary input to infer a P-cell's contribution to behavior

Complex spikes are rare events, occurring at around once per second. They briefly and completely suppress the SS rates, which induce downstream effects on the cerebellar nuclei (48), potentially producing movements (23). However, the CS rates are modulated to encode the direction of sensory prediction errors (8, 18, 49, 50). Thus, it was critical to test that the kinematic effects that we measured following a CS-induced SS suppression were not a consequence of a feedback response to kinematic deviations that occurred before the CS.

We did this by comparing triplets of temporally adjacent licks, finding that while the tongue trajectory preceding the CS event remained within the chance error bounds, the trajectory that followed

were robustly different than chance. Notably, the downstream effects of the CS-induced suppression were in the same direction, i.e., extension of the tongue, regardless of whether the CS events occurred during protraction or retraction. This is notable because the CS rates were maximum during protraction, and minimum during retraction, yet their downstream effects were the same: pull the tongue outwards.

The idea that the olivary input may affect ongoing movements was noted by Ebner and colleagues during reaching movements (24), and subsequently observed during saccades (25). For example, during saccades following the CS-induced SS suppression, the eyes are pulled in the CS-on direction of the P-cell. This is consistent with the fact that optogenetic increase in the SS rates suppresses the cerebellar nuclei (51), pulling the eyes approximately in direction CS+180 (52). While relying on the stochasticity in the olivary input has the disadvantage of lacking a causal manipulation, it has the advantage of relating the kinematic effects to CS-on properties of each P-cell, something that would not be possible with large scale optogenetic stimulation.

A long SS pause had the same effect on behavior as a CS-induced SS suppression

A CS is followed by SS suppression, but pauses in SS production can also occur because of other reasons, including inhibitory input from the molecular layer interneurons (53). Here we found that the effects of CS-induced SS suppression on tongue kinematics were largely the same as SS pauses that were not due to arrival of a CS. In both cases, the result was a force that pulled the tongue outwards and bent it away from the midline. This implies that the downstream effects on kinematics were not due to arrival of an input from the inferior olive, but rather the suppression or pausing of SS production in the parent P-cell.

A cortico-cerebellar network for control of the tongue

As the subject prepared to initiate a licking bout, the P-cells exhibited a ramping activity. Previous reports have noted that during this period there is ramping activity in the tongue regions of the motor cortex (54), as well as in the fastigial (54) and the dentate nuclei (55). Inhibiting the motor cortex in mice prevents both the onset and the termination of the licking bout (56), suggesting that both are active processes that are cortically mediated. Inhibiting the fastigial during the ramping period disrupts planning of the movement and removes the direction selectivity that the motor cortical cells exhibit (54), while inhibiting the dentate disrupts the ramping activity in the motor cortex (55). Similarly, exciting the P-cells in the vermis abolishes the ramping activity in motor cortex during the delay period of a decision-making task (57). This implies that as one prepares to initiate a movement, the rising activity in the motor cortex is controlled via a loop through the cerebellum.

Our results here suggest that once the tongue movement begins, there is a specific role for the cerebellum in producing forces that would stop the protraction, especially if that movement is reward relevant. We speculate that the cerebellum is informed via the mossy fibers of two kinds of information: the location of the target, and a copy of the ongoing motor commands (26). The function of this region of the cerebellum may be to use the copy of the motor commands to predict when the tongue is about to reach the target and aid in production of commands that would stop the outward movement.

Medial and lateral parts of the cerebellum may contribute to different aspects of tongue movements

In humans, the tongue region of the cerebellum extends from lobule VI in the vermis laterally to the hemispheres (58–60). Dysarthria is principally associated with damage in the paravermal regions of the

cerebellum (61). In macaques, stimulation of the fastigial nucleus moves the tongue in the protraction-retraction axis, while stimulation of the dentate nucleus moves it in the lateral-medial axis (10). In mice, activation of the P-cells in the lateral regions of lobule VI and VII during protraction bends the tongue toward the ipsilateral side (62). When we consider these results together with our observations here, what emerges is the conjecture that the P-cells in the vermis are important for control of protraction/retraction, but the P-cells in the paravermis and hemispheres have a different role, possibly in controlling how the tongue bends.

Toward a general model of how the cerebellum controls movements

Like the SS rates, the CS rates peaked near protraction peak velocity, then fell below baseline before the onset of retraction. Thus, for both ipsilateral and contralateral movements, the "CS-on action" across the P-cells was protraction, while the "CS-off action" was retraction. Notably, the downstream effects of SS suppression were to extend the tongue. As a result, there was a correspondence between the vector that described the CS-on action, and the vector that described the effects of SS suppression. This fact is notable because the same principle holds for P-cells in the oculomotor region of the cerebellum during saccadic eye movements (25, 26): the olivary input to an oculomotor P-cell is most active when a saccade is planned in direction CS-on, and the downstream effects of that P-cell's SS suppression is to pull the eyes also in direction CS-on. Thus, for both eye movements and tongue movements, the olivary input provides a vector based coordinate system (26) with which one might estimate the downstream contributions of a P-cell to control of that movement (63, 64).

The key theoretical idea is that the inferior olive organizes the cerebellum so that the P-cells are placed in competition with each other: for every P-cell that has a particular CS-on, effecting movements along a particular potent vector, there is another that prefers the opposite vector (26). Unfortunately, here we could not apply this theory to organize the P-cells into antagonist populations because nearly all the cells in our database had a CS response that peaked during protraction. However, our theory (26) predicts that there should be P-cells whose climbing fiber inputs prefer retraction. In these P-cells the SS suppression should pull the tongue inward. If these P-cells exist, then their SS pattern would be antagonistic to the SS pattern of the P-cells we found here, resulting in a population response in which P-cells would compete with each other, perhaps producing a sum of activity that is a burst-pause pattern, inhibiting then disinhibiting the nucleus as the tongue approaches the target.

Lingual dysfunction accompanies a host of symptoms, including vocal muscle dystonia (65), problems in swallowing (66), and dysarthria (2, 61, 67), all of which share a link to the cerebellum. Rehabilitation or cures for these symptoms will require a much better understanding of how the cerebellum contributes to control and learning of tongue movements. Marmosets are exceptionally skilled in shaping and twisting their tongue, using it almost like a finger. This makes them an attractive new model to study the neural control of a body part that is essential for our existence.

Methods

Data were collected from three marmosets, *Callithrix Jacchus*, 2 male and 1 female, 350-370 g, subjects 125D (Mirza), 59D (Ramon), and 132F (Charlie), during a 3.5-year period. The marmosets were born and raised in a colony that Prof. Xiaoqin Wang has maintained at the Johns Hopkins School of Medicine since 1996. The procedures on the marmosets were approved by the Johns Hopkins University Animal Care and Use Committee in compliance with the guidelines of the United States National Institutes of Health.

Data acquisition

Following recovery from head-post implantation surgery, the animals were trained to make saccades to visual targets and rewarded with a mixture of applesauce and lab diet (16). Visual targets were presented on an LCD screen. Binocular eye movements were tracked at 1000 Hz using EyeLink in subject R and M, and 2000 Hz using VPIX in subject C. Tongue movements were tracked with a 522 frame/sec Sony IMX287 FLIR camera, with frames captured at 100 Hz.

We performed MRI and CT imaging on each animal and used the imaging data to design an alignment system that defined trajectories from the burr hole to various locations in the cerebellar vermis (16), including points in lobule VI and VII. We used a piezoelectric, high precision microdrive (0.5 micron resolution) with an integrated absolute encoder (M3-LA-3.4-15 Linear smart stage, New Scale Technologies) to advance the electrode.

We recorded from lobules VI and VII of the cerebellum (Fig. 1C) using quartz insulated 4 fiber (tetrode) or 7 fiber (heptode) metal core (platinum/tungsten 95/05) electrodes (Thomas Recording), and 64 channel checkerboard or linear high density silicon probes (M1 and M2 probes, Cambridge Neurotech). We connected each electrode to a 32 or 64 channel head stage amplifier and digitizer (RHD2132 and RHD2164, Intan Technologies, USA), and then connected the head stage to a communication system (RHD2000 Evaluation Board, Intan Technologies, USA). Data were sampled at 30 kHz and band-pass filtered (2.5 - 7.6 kHz).

The silicon probes arrived with a polymer coating on the contacts that degraded with each insertion into the brain (68). This degradation increased the impedance of the electrodes and dramatically reduced the ability of the probe to isolate neurons. We found it essential to rejuvenate the silicon probes by stripping and then re-depositing the polymer coating after every 3-4 insertions into the brain (68).

Behavioral protocol

Each trial began with fixation of a center target after which a primary target appeared at one of 8 randomly selected directions at a distance of 5-6.5 deg. As the subject made a saccade to this primary target, that target was erased, and a secondary target was presented at a distance of 2-2.5 deg, also at one of 8 randomly selected directions. The subject was rewarded if following the primary saccade, it made a corrective saccade to the secondary target, landed within 1.5 deg radius of the target center, and maintained fixation for at least 200 ms. The food was provided via two small tubes (4.4 mm diameter), one to the left and the other to the right of the animal, positioned at 90° with respect to the mouth. A successful trial produced a food increment in one of the tubes and would continue to do so for 50-300 consecutive trials, then switch to the other tube. Because the food increment was small, the subjects naturally chose to work for a few consecutive trials, tracking the visual targets and allowing the food to accumulate, then stopped tracking and harvested the food via a licking bout. The subjects did not work while harvesting, and often fixated the tube. As a result, the behavior consisted of a work

period of targeted saccades, followed by a harvest period of targeted licking, repeated hundreds of times per session.

We measured eye movements during all phases of the task, including the bouts of licking. The monkeys tended to fixate the tube while licking. We analyzed tongue movements using DeepLabCut (69). Our network was trained on 89 video recordings of each subject with 15-25 frames extracted and labeled from each recording. The network was built on the ResNet-152 pre-trained model, and then trained over 1.03x106 iterations with a batch size of 8, using a GeForce GTX 1080Ti graphics processing unit. A Kalman filter was further applied to improve quality and smoothness of the tracking, and the output was analyzed in MATLAB to quantify lick events and kinematics. We tracked the tongue tip and the edge of the food in the tube, along with control locations (nose position and tube edges). We tracked all licks, regardless of whether they were aimed toward a tube, or not. Food-tube licks were further differentiated based on whether they aimed to enter the tube (inner-tube licks) or hit the outer edge of the tube (outer-edge licks). If any of these licks successfully contacted the food, we labeled that lick as a success (otherwise, an unsuccessful lick).

Tracking the tongue

The following videos provide examples of the various types of licks, along with the kinematic measures that we used to track each movement: **videos 1-10**. Licks were categorized based on heuristics that considered the position of the tongue relative to the tube opening and the food. We tracked 4 regions of the tongue consisting of the tip, the midpoint, and the left and right edges. The midpoint was computed based on the distance between the tip marker and the opening of the mouth, while the left and right edges were computed based on the mid-distance between the tip and midpoint, positioned at max laterality. Furthermore, we tracked the left and right edges of the opening of each reward tube as well as the densest edge of the food contained within.

Licks were labeled as reward seeking when the region of the tongue within the marker overlapped with the edge of the tube coordinates. Alternatively, licks were labeled as grooming when no overlap occurred. Reward seeking licks were further labeled into subcategories, consisting of inner-tube and outer-tube licks. Inner-tube labels were assigned when the tip, left, and right tongue markers remained within the bounds of the tube edge markers. Outer-tube labels were assigned when at least one marker exited the tube boundaries, conditioned on the tip having remained within at least 5 mm of the tube opening.

Additional labels were assigned to each reward seeking licks, categorizing them as either successful or unsuccessful licks. In all cases, overlap with food dictated these labels. Thus, to call a given lick an unsuccessful lick, the position of the food within the tube, relative to the tongue, was considered. For example, consider a scenario in which the food is depleted, requiring an inner-tube lick to scoop out the remaining bolus. If the lick entered the tube and thus touched the food, it was considered a success. If it did not enter the tube and thus did not touch the food, it was considered an unsuccessful lick.

Neurophysiological analysis

We used OpenEphys (70) for electrophysiology data acquisition, and then used P-sort (28) to identify the simple and complex spikes in the heptodes and tetrodes recordings. We used Kilosort and Phi (71) to identify the spikes for the silicon probes. Simple and complex spike instantaneous firing rate were calculated from peri-event time histograms with 1 ms bin size. We used a Savitzky–Golay filter (2nd order, 31 datapoints) to smooth the traces for visualization purposes.

Many P-cells in lobules VI and VII of the vermis were modulated during licking as well as during saccades. Our data here were selected from recordings that isolated P-cells with strong tongue related activity. The strength of behavioral modulation for each P-cell during saccades and licks was quantified using a z-score (Supplementary Fig. S2B). This z-score was calculated for each behavior via the range of the P-cell's average stimulus-aligned response divided by the standard deviation of this range, as computed across 2,000 permuted responses. Range was defined as the maximum change in firing rate from pre-behavior to post-behavior for a given response. This approach relies on the notion that if a cell is responsive to a given stimulus, it will exhibit both a strong response (high range) and a consistent response (low standard deviation of range values). Consistent with earlier work (9), the threshold for significant modulation during licking was set at a z-score of 3.

CS baseline firing rates were computed by dividing the total number of spikes by the duration of the entire recording. SS baseline firing rates were computed using two different methods depending on the analysis. For bout related responses, baseline was defined as the average firing rate in a 300 ms window preceding bout onset by 700 ms, i.e. during the [-1000 to -700] ms period. However, to analyze the activities during individual licks, because the rates were not stationary but gradually changing from the first to the last lick in the bout, baseline SS rates were computed using the average firing within a sliding window of 2 seconds, consisting of 5-6 licks.

To explore how the SS rates changed with the kinematic parameters of the lingual movements, we visualized the firing rates as a function of tongue endpoint position. The firing rates of each P-cell during maximal tongue velocity were computed on a trial-by-trial basis and associated with the spatial coordinates corresponding to the endpoint of that trial's lick. Single trial spike data was smoothed with a Savitzky-Golay filter. Spatial coordinates were standardized across animals such that all contralateral licks appear to the left and all ipsilateral licks to the right. A 100x50 grid was mapped onto the full range of tongue endpoint values, and the population firing rates at each point were estimated using a natural neighbor interpolation, effectively weighing contributions of neighboring firing rate values based on proximity. The interpolated surface was then smoothed with a 2-D Gaussian filter to produce a continuous heatmap. To ensure population-level robustness of firing rate values, a cell coverage mask was then applied over the heatmap, removing any grid points that did not have at least 75% of the available PCs (118/157 SS cells).

Computing the kinematic effects of CS-induced SS suppression

For each P-cell we considered triplets of tube-directed licks $\{n-1, n, n+1\}$, where all three licks were of the same type, i.e., contacted the same part of the tube (edge or inner). We then selected the subset of triplets in which there was a CS at only a single period in lick n, but no CS during any period in the two neighboring licks $n-1$, and $n+1$. We then compared tongue trajectories between the lick that had a CS with the two neighboring licks, i.e., $n-(n-1)$ and $n-(n+1)$.

Computing the kinematic effects of SS pauses

To assess if the perturbation of tongue movements was a consequence unique to the presence of a CS, or rather the suppression of SSs, we considered the effect of SS pauses on the tongue trajectory during licking, i.e., long ISI events that were not preceded by a CS.

For each P-cell, we selected the subset of all licks of the same type towards the same direction in which no CSs occurred at any point in the movement. Let us call these the NoCS licks. Working only with the NoCS licks, we sought to identify licks in which during a phase of interest (e.g., protraction

deceleration), there was a long pause in the SS production. However, we had to ensure that if there was a long pause in one phase of the lick, it did not also occur in other phases of the same lick. That is, like the CS analysis, to be eligible for this analysis a long SS pause had to occur only once during the lick.

There were 4 phases for each lick (protraction acceleration and deceleration, retraction acceleration and deceleration), i.e., $p = 1, \ldots, 4$. For each lick *n,* during each phase *p,* we found the duration of the longest ISI that originated in that phase (regardless of whether it extended into the next phase) and labeled it as $t_n^{(p)}$. Next, for each phase, we found the distribution of $t_n^{(p)}$. Licks with zero SSs during the given phase were excluded from this distribution.

For example, suppose we were interested in labeling the licks in which during phase 1 there was a long pause. A lick with a long pause in phase 1 could not have also had a long pause in another phase of that same lick. We found the distribution of $t_n^{(2)}$, the distribution of $t_n^{(3)}$, and the distribution of $t_n^{(4)}$, and then for each phase selected the top quartile (25% longest ISIs). We removed the licks with a long pause in phase 2-4 for consideration. From among the remaining licks, we formed the distribution of $t_n^{(1)}$, found the top quartile, and labeled those as having a long pause during phase 1. We labeled the remaining 75% of licks in this population as not having a long pause during this phase.

We selected the subset of triplets in which lick *n* had a long SS pause in only one phase of the movement, but no SS pause occurred in any phase of the two neighboring licks. We then compared tongue trajectories between the lick that had a pause with the two neighboring licks. Traces were averaged within directions and then across directions for each cell.

Computing the effects of trajectory error on climbing fiber activity

Roughly 15% of the licks failed to enter the tube and did not touch the food (Supplementary Fig. S5). To visualize the complex spike patterns as a function of the spatial location of the tongue, we began with computing $p(x(t-25)|CS)$, i.e., given that a CS occurred at time t, the likelihood of the tongue's tip location x at time $t-25$ms. We did this by averaging the position of the tip of the tongue during the 50ms period before the CS event. We separated the licks into successful licks (tongue entered the tube and touched the food) and unsuccessful licks (tongue touched the tube but neither entered it nor touched the food). The result was the likelihood $p(x_s|CS)$ for the successful licks and $p(x_u|CS)$ for the unsuccessful licks.

We next computed the marginal probability density $p(x)$ for each lick type, the prior $\Pr(CS)$ (from the average CS rate during a lick of that type, using a 50ms time bin), and then the ratio of the probabilities $p(x|CS)\Pr(CS)/p(x)$. Finally, we computed the error-induced spatial pattern of complex spikes by subtracting this ratio for the successful licks from the same ratio for the unsuccessful licks. To reduce the noise associated with the far edges of each probability density function, for each ratio we considered values that were in the 95% quantile of its distribution.

Statistical analysis

In order to compare the measured effect of SS suppression on tongue trajectory with what would be expected to happen simply due to chance, we computed the bounds for the null hypothesis. To do so, we used bootstrapping to compute 95% confidence intervals. We shuffled the assignment of CS tags from the lick in which it had occurred to a randomly assigned lick of the same type. Using this pseudo-data, we then selected triplets of consecutive tube-directed licks and computed trajectory differences among neighboring licks, averaging $n-(n-1)$ and $n-(n+1)$. We computed this expected value for

each cell, computed a mean across all the cells, and then repeated the shuffling 30 times to compute 95% confidence intervals.

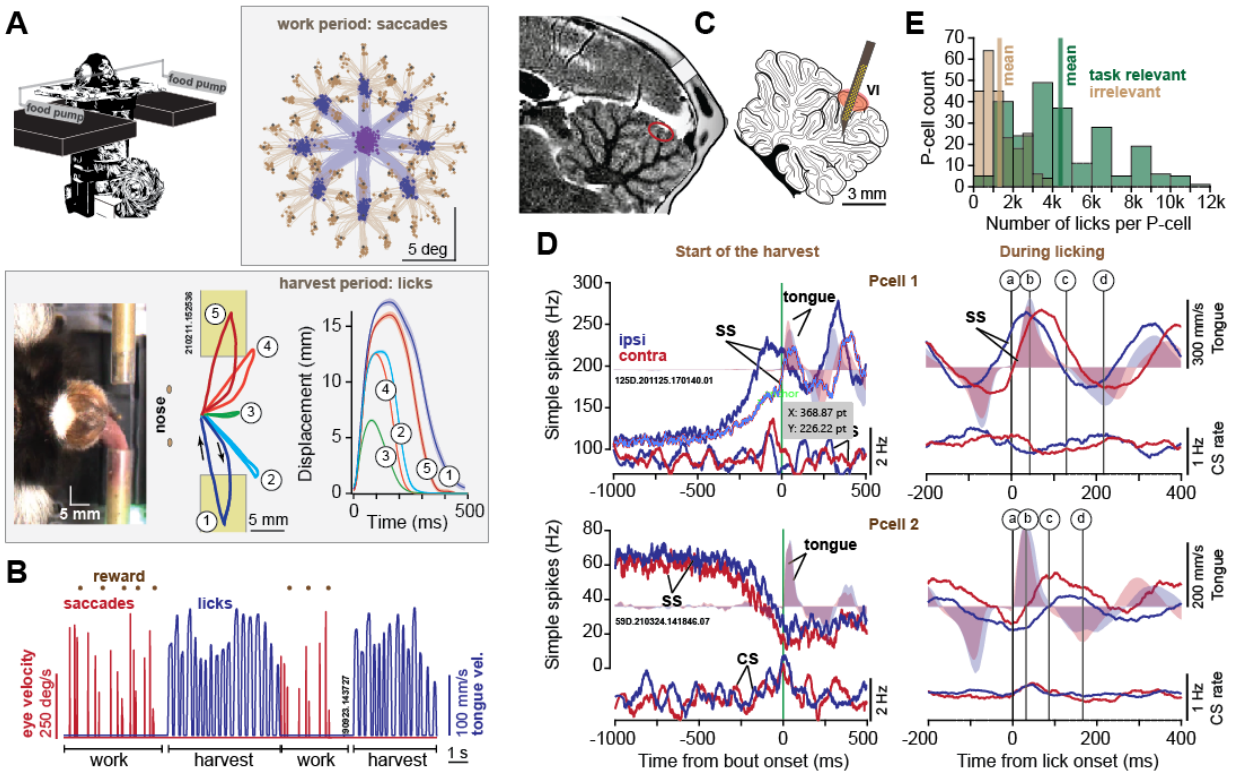

Figure 1. Marmosets produced dexterous tongue movements during recordings from the cerebellar vermis. A. Subjects made saccades to visual targets and received a small amount of food as reward via one of two tubes placed obliquely to the mouth. In the task relevant licks, they directed their tongue to the edge of the tube to harvest food near the tip (trajectories 2, 4), or inside the tube to harvest food that was deeper (trajectories 1, 5). In task irrelevant licks, they groomed their face (trajectory 3). **B**. Subjects chose to work for consecutive trials, making saccades and allowing the food to accumulate, then harvested their cache in bouts of licking. **C**. We employed silicon probes to record from lobule VI and VII of the vermis. **D**. Simple and complex spikes (SS, CS) of two Purkinje cells, aligned to bout onset and lick onset. A single lick was divided into acceleration period of protraction (a-b), deceleration period of protraction (b-c), and acceleration period of retraction (c-d). Filled color regions indicate tongue velocity. **E**. The number of task relevant (tube directed) and task irrelevant (grooming) licks recorded per neuron.

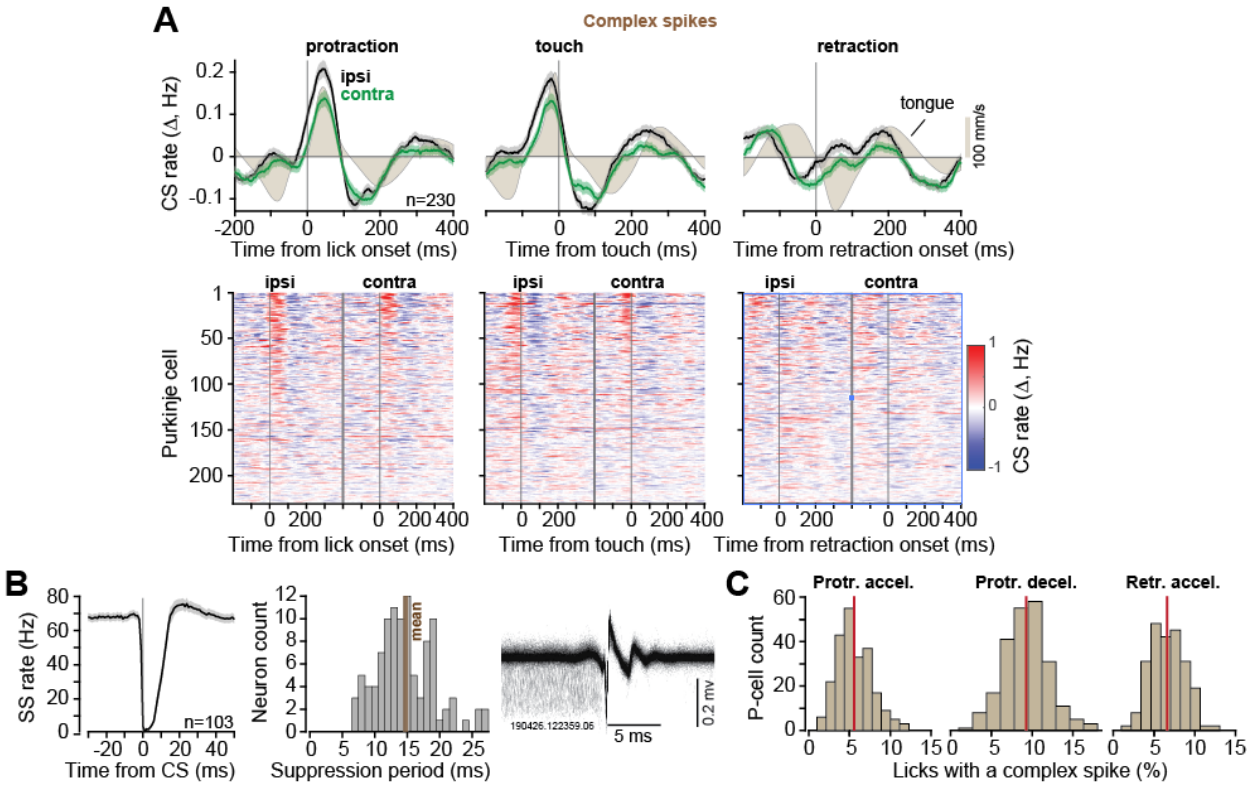

Figure 2. CS rates increased with protraction and decreased with retraction. A. CS rates across the population aligned to protraction onset, touch of the tube, and retraction onset. The second row shows CS activity across all P-cells, aligned to protraction, touch, and retraction, sorted based on CS rate for ipsilateral licks at lick onset. **B**. CS-induced SS suppression, averaged across the subset of P-cells for which both the CS and the SS were isolated (left). Examples of SS and CS waveforms for a single P-cell are shown at right. **C**. Percentage of licks with a CS during a specific period of time for each neuron. Vertical line indicates mean. Error bars are SEM.

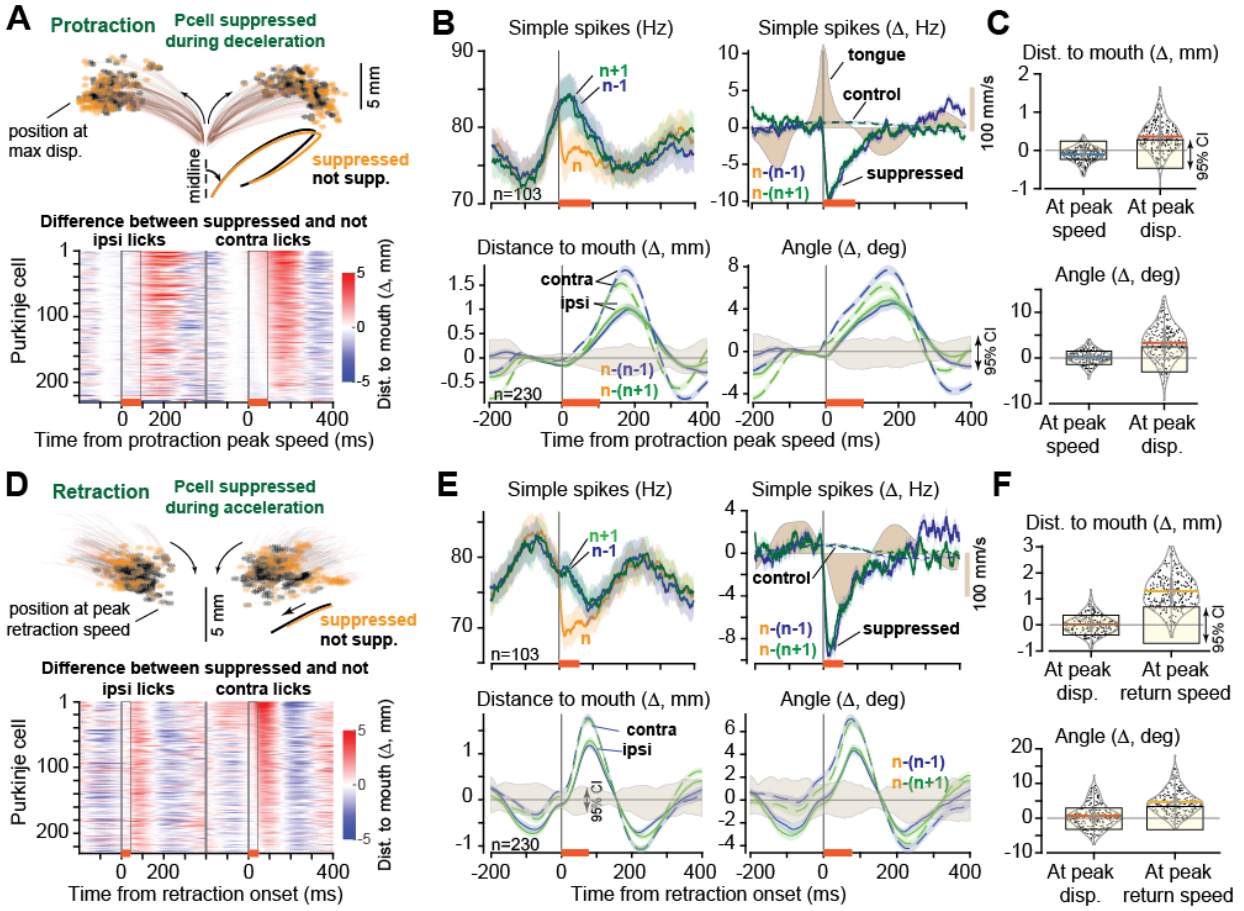

Figure 3. CS-induced SS suppression produced hypermetria during protraction and slowing during retraction. A. Suppression took place during the deceleration period of protraction. Traces show average tongue trajectory during this period of protraction for each P-cell during SS suppressed and control licks. Ipsilateral licks are shown to the left and contralateral to the right. Heatmap quantifies change in endpoint trajectory between suppressed and control licks for each cell. Period of suppression is indicated by the orange bar at the bottom of the heatmap. **B**. Top row: SS rates for licks { , , }, where only lick experienced a CS. Filled color curves indicate tongue velocity. Second row: trajectory of the tongue in lick as compared to its two temporally neighboring licks. Trajectory is measured via distance from tip of the tongue to the mouth and angle of the tip with respect to midline. The filled region is 95% CI. **C**. Distance to mouth and angle in lick as compared to neighboring licks. Shaded region is 95% CI. **D-F**. Suppression during the acceleration period of retraction induced slowing. Same format as in parts A-C. In part F, tongue trajectory (displacement and angle) is similar before SS suppression (at peak displacement) but diverges after the suppression at peak return speed. Error bars are SEM.

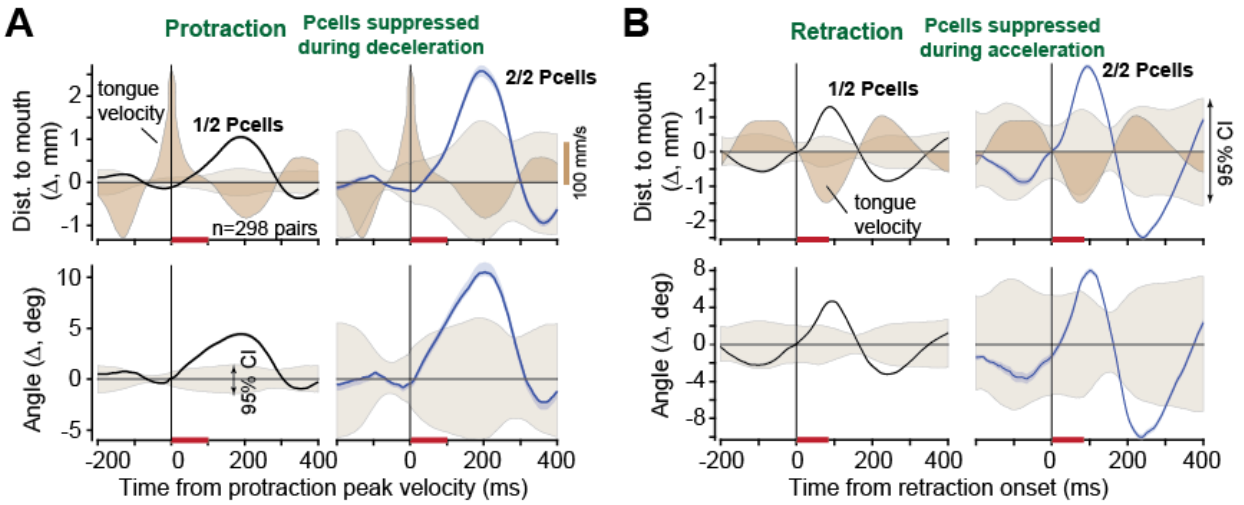

Figure 4. Suppression of multiple P-cells scaled the perturbation to the tongue. **A**. Kinematic effects of CS-induced P-cell suppression during the deceleration period of protraction (orange bar). Traces show change in tongue trajectory in suppressed licks vs. control licks, measured via distance from tip of the tongue to the mouth and angle of the tip with respect to midline. Gray shaded region is 95% CI. Brown filled region is tongue velocity. **B**. Kinematic effects of P-cell suppression during the acceleration period of retraction (orange bar). Same format as in part A. Error bars are SEM.

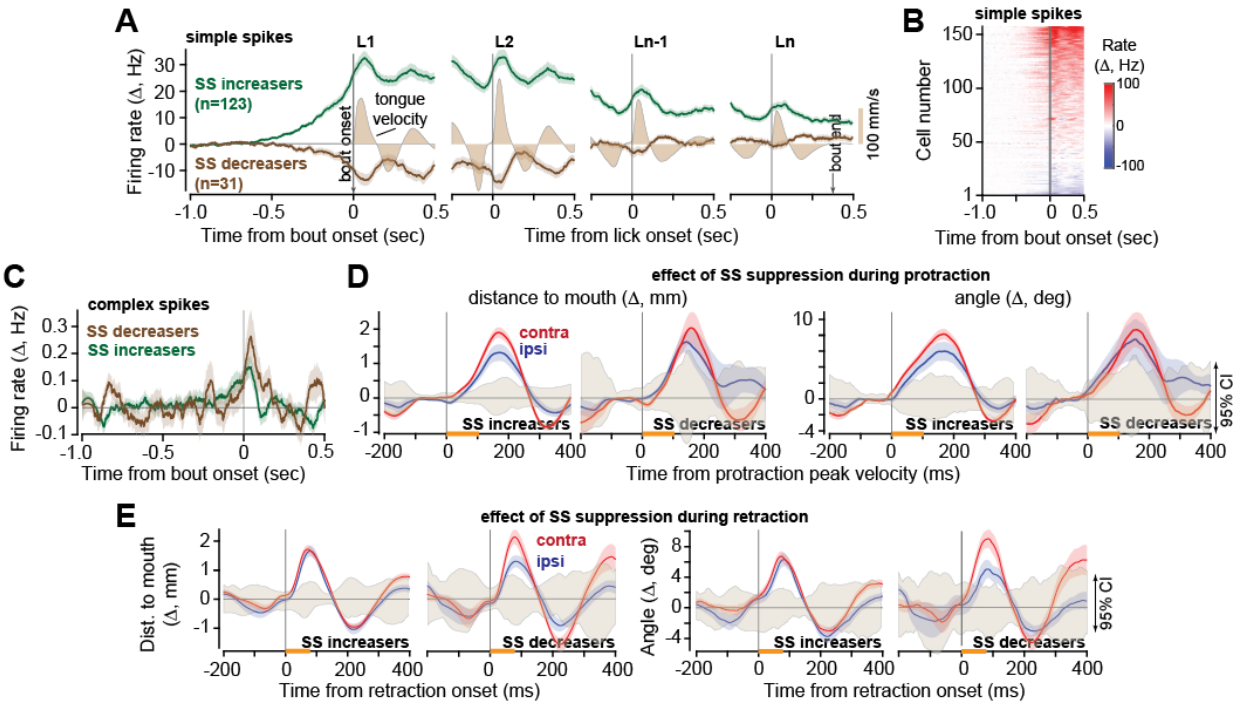

Figure 5. Effects of SS suppression remained consistent across P-cells regardless of SS modulation. A. Some of the P-cells exhibited an increase in SS rates before bout onset, while others exhibited a decrease. L1 is first lick in the bout, Ln is last lick. **B**. Change in SS rates for all P-cells (with respect to baseline), aligned to bout onset. **C**. Change in complex spike rate for the same P-cells. **D**. Kinematic effects of SS suppression during the deceleration period of protraction (orange bar). Shaded region is 95% CI. **E.** Kinematic effects of SS suppression during the acceleration period of retraction (orange bar). Error bars are SEM.

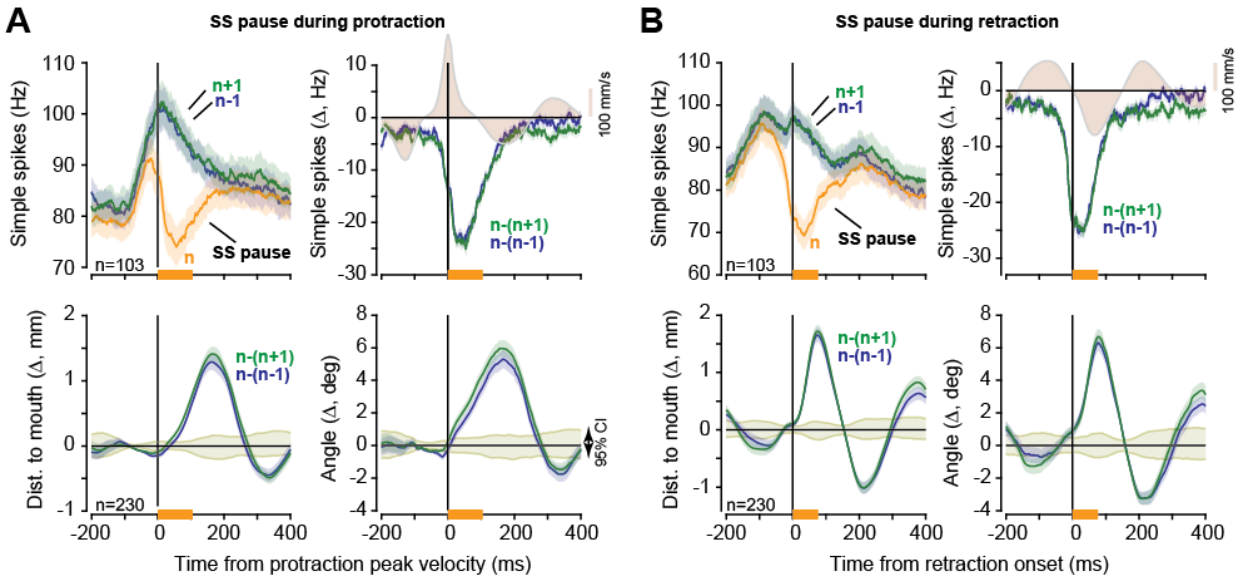

Figure 6. SS pause without a CS was sufficient to produce hypermetria and bending of the tongue. A. We selected licks that did not have a CS but nevertheless experienced a long SS pause during the protraction deceleration period. Top row: SS rates for lick n that experienced a pause and licks n-1 and n+1 that did not. Bottom row: change in lick kinematics following the SS pause. **B.** Same as in part A, but for licks that experience a long SS pause during the retraction acceleration period.

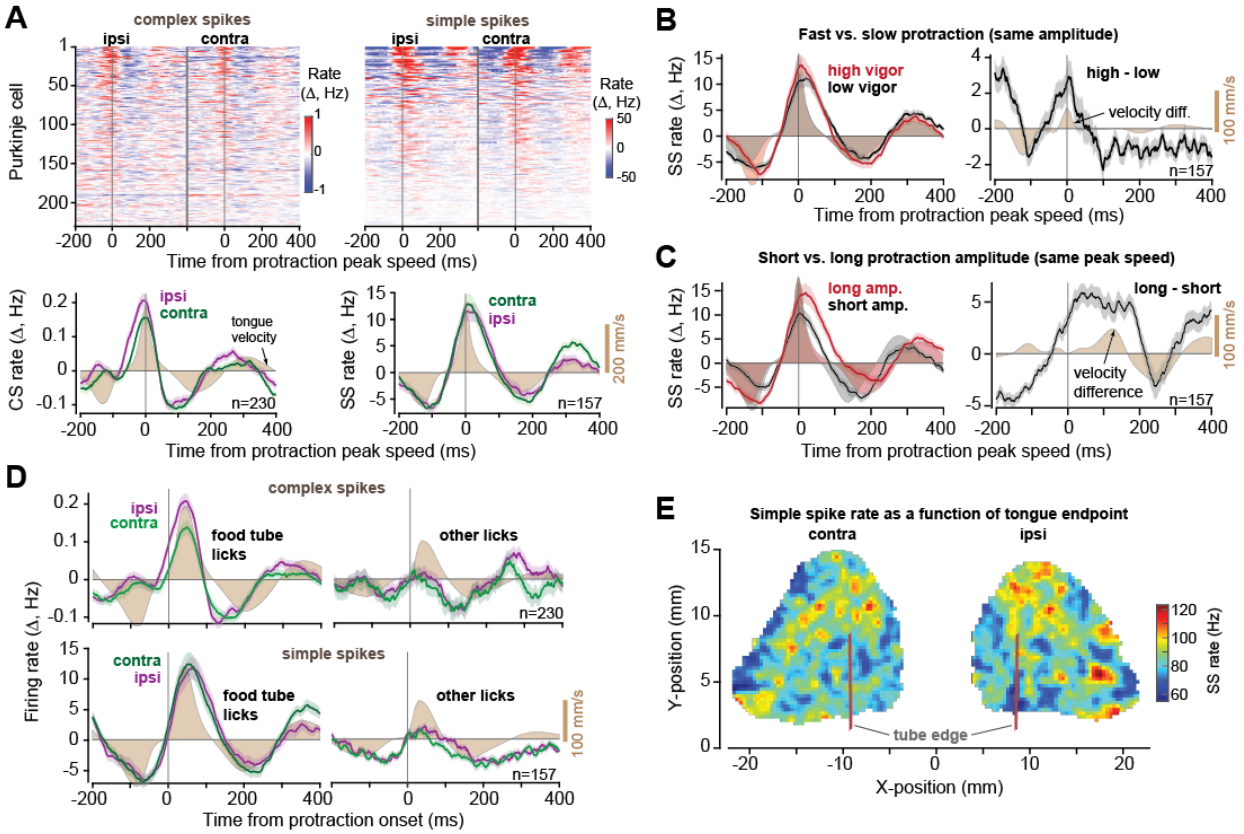

Figure 7. SS rates peaked at deceleration onset, but only if the lick was task relevant. A. CS and SS rates across the population aligned to protraction peek speed. **B**. Left figure shows SS rates for high and low vigor protractions (same amplitude). Right figure shows within cell difference in SS rates for high minus low vigor licks. Shaded regions are tongue speed for high and low vigor licks (left) and change in speed (right). **C**. Left figure shows SS rates for protractions that had long or short amplitudes but the same peak speed. Right figure shows within cell differences. Shaded regions are tongue speed for long and short licks (left) and change in speed (right). **D**. Modulation of CS and SS rates during task relevant and task irrelevant licks. Light shaded region is tongue speed for task relevant movements, while dark shaded region is tongue speed for task irrelevant movements. **E**. SS rates at peak velocity as a function of tongue position at maximum displacement. Error bars are SEM.

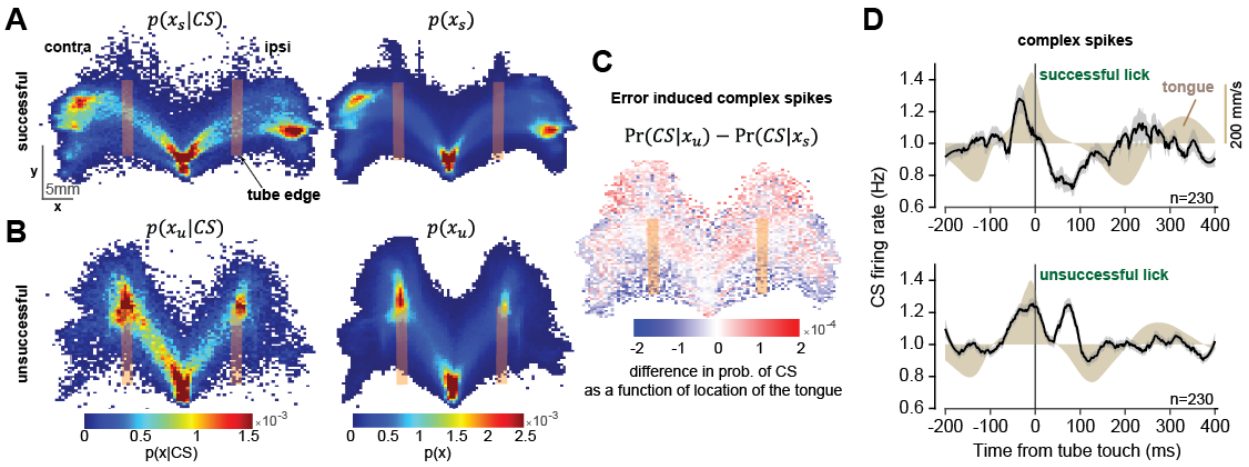

Figure 8. Error in the tongue's trajectory induced complex spikes. **A**. Successful licks: tongue entered the tube and touched the food. Left subfigure shows the spatial likelihood at 25 ms before the CS event. Right subfigure shows the probability of being at a spatial location. **B**. Unsuccessful licks: tongue did not enter the tube and collided with the tube's edge. **C**. Spatial pattern of the error induced complex spikes (difference between unsuccessful and successful licks in the posterior probability of CS as a function of the tongue's location). **D**. Complex spike rates following tube touch for successful and unsuccessful licks.

Video 1. A sequence of 5 licks to the right tube. The top two plots show trajectory of the tip of the tongue and a geometric representation of four markers on the tongue. The 2nd row shows the displacement, velocity, and angle of the tongue as a function of time. The 3rd row shows the distance of the tip of the tongue to the food in the left and the right tube. The first 3 licks are successful and enter the tube and contact the food. In the 4th and 5th licks the tongue fails to enter the tube and do not contact the food. These licks are unsuccessful and are analyzed in Fig. 8.

Video 2. Example of a grooming lick. These licks aim to clean the regions around the mouth and are not aimed toward the food tubes.

Video 3. Example of a grooming lick.

Video 4. Example of a bout of grooming licks.

Video 5. Example of an outer-tube lick. The food has accumulated beyond the edge of the tube and the subject begins the bout by licking the food near the edge.

Video 6. Example of an inner-tube lick. The food is deep inside the tube and the subject enters the tube and scoops the food out.

Video 7. Example of an inner-tube lick.

Video 8. Unsuccessful lick. The food is inside the tube, but the lick fails to enter it and instead goes under the tube.

Video 9. Unsuccessful lick. The food is inside the tube, but the lick fails to enter it and instead goes to the outer edge.

Video 10. Unsuccessful lick. The food is inside the tube, but the lick fails to enter it and instead collides with the edge.

1. W. J. M. Levelt, *Speaking: From Intention to Articulation* (MIT Press, 1993).

2. M. Gentil, EMG analysis of speech production of patients with Friedreich disease. *Clin. Linguist. Phon.* **4**, 107–120 (1990).

3. T. Vilis, J. Hore, Central neural mechanisms contributing to cerebellar tremor produced by limb perturbations. *J.Neurophysiol.* **43**, 279–291 (1980).

4. F. Yu, Q. J. Jiang, X. Y. Sun, R. W. Zhang, A new case of complete primary cerebellar agenesis: clinical and imaging findings in a living patient. *Brain* **138**, e353 (2015).

5. J. L. Bryant, J. D. Boughter, S. Gong, M. S. LeDoux, D. H. Heck, Cerebellar cortical output encodes temporal aspects of rhythmic licking movements and is necessary for normal licking frequency. *Eur. J. Neurosci.* **32**, 41–52 (2010).

6. M. A. Gaffield, J. M. Christie, Movement Rate Is Encoded and Influenced by Widespread, Coherent Activity of Cerebellar Molecular Layer Interneurons. *J.Neurosci.* **37**, 4751–4765 (2017).

7. M. I. Becker, A. L. Person, Cerebellar Control of Reach Kinematics for Endpoint Precision. *Neuron* **103**, 335–348 (2019).

8. J. S. Pi, *et al.*, The olivary input to the cerebellum dissociates sensory events from movement plans. *Proc. Natl. Acad. Sci.* **121**, e2318849121 (2024).

9. E. Sedaghat-Nejad, J. S. Pi, P. Hage, M. A. Fakharian, R. Shadmehr, Synchronous spiking of cerebellar Purkinje cells during control of movements. *Proc. Natl. Acad. Sci.* **119**, e2118954119 (2022).

10. J. P. Bowman, L. D. Aldes, Organization of the cerebellar tongue representation in the monkey. *Exp. Brain Res.* **39**, 249–259 (1980).

11. L. D. Aldes, J. P. Bowman, Representation of the tongue in the cerebellar nuclei of the monkey. *Exp. Neurol.* **64**, 202–215 (1979).

12. M. A. Gaffield, B. A. Sauerbrei, J. M. Christie, Cerebellum encodes and influences the initiation, performance, and termination of discontinuous movements in mice. *eLife* **11**, e71464 (2022).

13. J. P. Welsh, E. J. Lang, I. Suglhara, R. Llinas, Dynamic organization of motor control within the olivocerebellar system. *Nature* **374**, 453–457 (1995).

14. K. Kubota, S. Hayama, Comparative anatomical and neurohistological observations on the tongues of pigmy and common marmosets. *Anat. Rec.* **150**, 473–485 (1964).

15. G. Oren, *et al.*, Vocal labeling of others by nonhuman primates. *Science* **385**, 996–1003 (2024).

16. E. Sedaghat-Nejad, *et al.*, Behavioral training of marmosets and electrophysiological recording from the cerebellum. *J.Neurophysiol.* **122**, 1502–1517 (2019).

17. P. Hage, *et al.*, Effort cost of harvest affects decisions and movement vigor of marmosets during foraging. *eLife* **12**, RP87238 (2023).

18. D. J. Herzfeld, Y. Kojima, R. Soetedjo, R. Shadmehr, Encoding of error and learning to correct that error by the Purkinje cells of the cerebellum. *Nat.Neurosci.* **21**, 736–743 (2018).

19. C. Ju, *et al.*, Neurons of the inferior olive respond to broad classes of sensory input while subject to homeostatic control. *J. Physiol.* **597**, 2483–2514 (2019).

20. S. Tsutsumi, *et al.*, Purkinje Cell Activity Determines the Timing of Sensory-Evoked Motor Initiation. *Cell Rep.* **33**, 108537 (2020).

21. Y. Sato, A. Miura, H. Fushiki, T. Kawasaki, Short-term modulation of cerebellar Purkinje cell activity after spontaneous climbing fiber input. *J. Neurophysiol.* **68**, 2051–2062 (1992).

22. W. T. Thach, Somatosensory receptive fields of single units in cat cerebellar cortex. *J.Neurophysiol.* **30**, 675–696 (1967).

23. V. Romano, *et al.*, Olivocerebellar control of movement symmetry. *Curr. Biol.* **32**, 654-670.e4 (2022).

24. M. L. Streng, L. S. Popa, T. J. Ebner, Climbing Fibers Control Purkinje Cell Representations of Behavior. *J.Neurosci.* **37**, 1997–2009 (2017).

25. S. Z. Muller, *et al.*, Complex spikes perturb movements and reveal the sensorimotor map of Purkinje cells. *Curr. Biol.* **33**, 4869-4879.e3 (2023).

26. M. A. Fakharian, A. M. Shoup, P. Hage, H. Y. Elseweifi, R. Shadmehr, A vector calculus for neural computation in the cerebellum. [Preprint] (2024). Available at: https://www.biorxiv.org/content/10.1101/2024.11.14.623565v1 [Accessed 18 November 2024].

27. A. Mathis, *et al.*, DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nat. Neurosci.* **21**, 1281–1289 (2018).

28. E. Sedaghat-Nejad, *et al.*, P-sort: an open-source software for cerebellar neurophysiology. *J. Neurophysiol.* **126**, 1055–1075 (2021).

29. T. R. Reppert, *et al.*, Movement vigor as a trait-like attribute of individuality. *J.Neurophysiol.* **120**, 741–757 (2018).

30. Y. Kojima, F. R. Robinson, R. Soetedjo, Cerebellar fastigial nucleus influence on ipsilateral abducens activity during saccades. *J.Neurophysiol.* **111**, 1553–1563 (2014).

31. C. Bourrelly, J. Quinet, L. Goffart, The caudal fastigial nucleus and the steering of saccades toward a moving visual target. *J. Neurophysiol.* **120**, 421–438 (2018).

32. C. W. Mohler, R. H. Wurtz, Organization of monkey superior colliculus: intermediate layer cells discharging before eye movements. *J.Neurophysiol.* **39**, 722–744 (1976).

33. D. M. Waitzman, T. P. Ma, L. M. Optican, R. H. Wurtz, Superior colliculus neurons mediate the dynamic characteristics of saccades. *J.Neurophysiol.* **66**, 1716–1737 (1991).

34. M. E. Goldberg, C. J. Bruce, Primate frontal eye fields. III. Maintenance of a spatially accurate saccade signal. *J.Neurophysiol.* **64**, 489–508 (1990).

35. B. S. Ito, Y. Gao, B. Kardon, J. H. Goldberg, A collicular map for touch-guided tongue control. *Nature* 1–9 (2025). https://doi.org/10.1038/s41586-024-08339-3.

36. T. M. Hoogland, J. R. De Gruijl, L. Witter, C. B. Canto, C. I. De Zeeuw, Role of Synchronous Activation of Cerebellar Purkinje Cell Ensembles in Multi-joint Movement Control. *Curr. Biol.* **25**, 1157–1165 (2015).

37. S. Kitazawa, T. Kimura, P. B. Yin, Cerebellar complex spikes encode both destinations and errors in arm movements. *Nature* **392**, 494 (1998).

38. A. L. Hewitt, L. S. Popa, T. J. Ebner, Changes in Purkinje cell simple spike encoding of reach kinematics during adaption to a mechanical perturbation. *J.Neurosci.* **35**, 1106–1124 (2015).

39. W. Heffley, *et al.*, Coordinated cerebellar climbing fiber activity signals learned sensorimotor predictions. *Nat. Neurosci.* **21**, 1431–1441 (2018).

40. M. J. Wagner, *et al.*, A neural circuit state change underlying skilled movements. *Cell* **184**, 3731-3747.e21 (2021).

41. D. J. Calame, M. I. Becker, A. L. Person, Cerebellar associative learning underlies skilled reach adaptation. *Nat. Neurosci.* 1–12 (2023). https://doi.org/10.1038/s41593-023-01347-y.

42. T. Ishikawa, *et al.*, Releasing dentate nucleus cells from Purkinje cell inhibition generates output from the cerebrocerebellum. *PLoS.One.* **9**, e108774 (2014).

43. Y. Kojima, R. Soetedjo, Change in sensitivity to visual error in superior colliculus during saccade adaptation. *Sci.Rep.* **7**, 9566 (2017).

44. Y. Kojima, R. Soetedjo, Elimination of the error signal in the superior colliculus impairs saccade motor learning. *Proc.Natl.Acad.Sci.U.S.A* **115**, E8987–E8995 (2018).

45. P. J. May, S. Warren, Y. Kojima, The superior colliculus projection upon the macaque inferior olive. *Brain Struct. Funct.* (2024). https://doi.org/10.1007/s00429-023-02743-7.

46. R. Soetedjo, Y. Kojima, A. F. Fuchs, How cerebellar motor learning keeps saccades accurate. *J. Neurophysiol.* **121**, 2153–2162 (2019).

47. A. Thomas, *et al.*, Superior colliculus bidirectionally modulates choice activity in frontal cortex. *Nat. Commun.* **14**, 7358 (2023).

48. T. Tang, T. A. Blenkinsop, E. J. Lang, Complex spike synchrony dependent modulation of rat deep cerebellar nuclear activity. *eLife* **8**, e40101 (2019).

936 49. J. F. Medina, S. G. Lisberger, Links from complex spikes to local plasticity and motor learning in the
937 cerebellum of awake-behaving monkeys. *Nat.Neurosci.* **11**, 1185–1192 (2008).

938 50. Y. Yang, S. G. Lisberger, Role of Plasticity at Different Sites across the Time Course of Cerebellar
939 Motor Learning. *J. Neurosci.* **34**, 7077–7090 (2014).

940 51. J. Chaumont, *et al.*, Clusters of cerebellar Purkinje cells control their afferent climbing fiber
941 discharge. *Proc.Natl.Acad.Sci.U.S.A.* **110**, 16223–16228 (2013).

942 52. R. Soetedjo, G. D. Horwitz, Closed-Loop Optogenetic Perturbation of Macaque Oculomotor
943 Cerebellum: Evidence for an Internal Saccade Model. *J. Neurosci.* **44** (2024).

944 53. M. Jelitai, P. Puggioni, T. Ishikawa, A. Rinaldi, I. Duguid, Dendritic excitation–inhibition balance
945 shapes cerebellar output during motor behaviour. *Nat. Commun.* **7**, 13722 (2016).

946 54. Z. Gao, *et al.*, A cortico-cerebellar loop for motor planning. *Nature* **563**, 113–116 (2018).

947 55. F. P. Chabrol, A. Blot, T. D. Mrsic-Flogel, Cerebellar Contribution to Preparatory Activity in Motor
948 Neocortex. *Neuron* **103**, 506–519 (2019).

949 56. D. Xu, *et al.*, Cortical processing of flexible and context-dependent sensorimotor sequences.
950 *Nature* **603**, 464–469 (2022).

951 57. J. Zhu, H. Hasanbegović, L. D. Liu, Z. Gao, N. Li, Activity map of a cortico-cerebellar loop underlying
952 motor planning. *Nat. Neurosci.* 1–13 (2023). https://doi.org/10.1038/s41593-023-01453-x.

953 58. J. Diedrichsen, E. Zotow, Surface-Based Display of Volume-Averaged Cerebellar Imaging Data. *PLOS*
954 *ONE* **10**, e0133402 (2015).

955 59. K. Grabski, *et al.*, Functional MRI assessment of orofacial articulators: Neural correlates of lip, jaw,
956 larynx, and tongue movements. *Hum. Brain Mapp.* **33**, 2306–2321 (2012).

957 60. W. Grodd, E. Hülsmann, M. Lotze, D. Wildgruber, M. Erb, Sensorimotor mapping of the human
958 cerebellum: fMRI evidence of somatotopic organization. *Hum. Brain Mapp.* **13**, 55–73 (2001).

959 61. P. P. Urban, *et al.*, Cerebellar Speech Representation: Lesion Topography in Dysarthria as Derived
960 From Cerebellar Ischemia and Functional Magnetic Resonance Imaging. *Arch. Neurol.* **60**, 965–972
961 (2003).

962 62. L. Bina, *et al.*, Cerebellar control of targeted tongue movements. [Preprint] (2024). Available at:
963 https://www.biorxiv.org/content/10.1101/2024.09.26.615128v1 [Accessed 6 November 2024].

964 63. T. J. H. Ruigrok, J. Voogd, Organization of projections from the inferior olive to the cerebellar nuclei
965 in the rat. *J. Comp. Neurol.* **426**, 209–228 (2000).

966 64. R. Apps, M. Garwicz, Anatomical and physiological foundations of cerebellar information
967 processing. *Nat.Rev.Neurosci.* **6**, 297–311 (2005).

65. A. Poologaindran, *et al.*, The effect of unilateral thalamic deep brain stimulation on the vocal dysfunction in a patient with spasmodic dysphonia: interrogating cerebellar and pallidal neural circuits. (2017). https://doi.org/10.3171/2016.10.JNS161025.

66. A. Sasegbon, S. Hamdy, The Role of the Cerebellum in Swallowing. *Dysphagia* **38**, 497–509 (2023).

67. R. Lechtenberg, S. Gilman, Speech disorders in cerebellar disease. *Ann. Neurol.* **3**, 285–290 (1978).

68. A. M. Shoup, *et al.*, Rejuvenating silicon probes for acute electrophysiology. [Preprint] (2024). Available at: https://www.biorxiv.org/content/10.1101/2024.02.20.581222v1 [Accessed 13 March 2024].

69. A. Mathis, *et al.*, DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nat. Neurosci.* **21**, 1281–1289 (2018).

70. J. H. Siegle, *et al.*, Open Ephys: an open-source, plugin-based platform for multichannel electrophysiology. *JNeural Eng* **14**, 045003 (2017).

71. M. Pachitariu, N. Steinmetz, S. Kadir, M. Carandini, H. K. D, Kilosort: realtime spike-sorting for extracellular electrophysiology with hundreds of channels. *bioRxiv* 061481 (2016). https://doi.org/10.1101/061481.