Cerebellar contributions to motor control and language comprehension: searching for common computational principles

Torgeir Moberget[1] and **Richard B. Ivry**[2]

[1]Norwegian Centre for Mental Disorders Research (NORMENT), KG Jebsen Centre for Psychosis Research, Division of Mental Health and Addiction, Oslo University Hospital, Norway

[2]Department of Psychology, and the Helen Wills Neuroscience Institute, University of California, Berkeley, Berkeley, California

Abstract

The past 25 years have seen the functional domain of the cerebellum extend beyond the realm of motor control, with considerable discussion of how this subcortical structure contributes to cognitive domains including attention, memory, and language. Drawing on evidence from neuroanatomy, physiology, neuropsychology, and computational work, sophisticated models have been developed to describe cerebellar function in sensorimotor control and learning. In contrast, mechanistic accounts of how the cerebellum contributes to cognition have remained elusive. Inspired by the homogeneous cerebellar microanatomy and a desire for parsimony, many researchers have sought to extend mechanistic ideas from motor control to cognition. One influential hypothesis centers on the idea that the cerebellum implements *internal models*, representations of the context-specific dynamics of an agent's interactions with the environment, enabling predictive control. We briefly review cerebellar anatomy and physiology, to review the internal model hypothesis as applied in the motor domain, before turning to extensions of these ideas in the linguistic domain, focusing on speech perception and semantic processing. While recent findings are consistent with this computational generalization, they also raise challenging questions regarding the nature of cerebellar learning, and may thus inspire revisions of our views on the role of the cerebellum in sensorimotor control.

Keywords

cerebellum; language; cognition; motor control; computational mechanisms; internal models

Introduction: the cerebellar cognitive revolution

Impaired coordination of movement is the most consistently observed symptom following lesions of the cerebellum, either induced in experimental animals[1] or arising because of

acute or degenerative pathology in humans.[2] Conversely, serious sensory deficits, such as blindness, or marked changes in cognition such as aphasia, are not associated with cerebellar lesions. Indeed, the neurologist Gordon Holmes, an early authority in the field, summed up his extensive years of clinical observations, writing "From numerous examinations I have no doubt, however, that in man even extensive lesions of the cerebellum involve no form of conscious sensation" (Ref. 2, p. 7). Physiological studies involving different methods to stimulate the cerebellum were also consistent with Holmes' conclusions: While stimulation of the cerebral cortex reliably evokes vivid sensations or perceptions,[3–5] stimulation of the cerebellum very rarely affects conscious experience.[6–8]

Thus, through the 20th century, textbooks of anatomy, neurology, and neuroscience limited the discussion of the cerebellum to chapters on motor control and sensorimotor learning.[9] Discussion of the complexities of the human mind were fixated on the cerebral cortex, resulting in what Parvizi has diagnosed as a corticocentric myopia in the cognitive neurosciences.[10] However, over the last few decades, the notion of cerebellar involvement in higher-level cognitive functions—including the quintessentially human faculty of language—has attracted considerable attention.[11–15]

The gauntlet for the "cerebellar cognitive revolution" was laid forth by Leiner *et al.*, who, in their seminal 1986 paper, proposed that the cerebellum contributes to cognitive as well as motor skills.[11] They went on a few years later to expand on this theme, outlining a potential role in language processing.[16] These conjectures were based on two main lines of reasoning. First, the evidence from comparative neuroanatomy suggested a rapid enlargement of the cerebellum over the course of vertebrate evolution,[11] a view that has been confirmed and expanded on by recent findings.[17,18] Indeed, throughout the evolution of apes, including humans, the rate of cerebellar expansion has even exceeded that of the neocortex,[18] leading the authors to propose that the change in cerebellar size "is likely to have underpinned the evolution of humans' advanced technological capacities, which in turn may have been a preadaptation for language."

Leiner *et al.* provided a second argument for a cerebellar role in cognition, turning here to emerging ideas on the computational capabilities of the cerebellum and how they might be used in networks that link the cerebellum and nonmotor areas of the cerebral cortex. We return to cerebellar microanatomy and physiology in the next section; here, we want to comment briefly on cerebrocerebellar connectivity. It has long been recognized that the cerebellum receives widespread and functionally diverse input from the cerebral cortex,[19,20] but in the classic view ascending cerebellar output was thought to target, via the thalamus, primary and secondary motor cortices.[19] This organization would be consistent with a view in which information from associative areas is primarily used to improve motor control. However, as shown in a series of elegant experiments using trans-synaptic tracers in primates, the output nuclei of the cerebellum also provide bisynpatic input to associative cortex, and thus provide an anatomical platform to influence nonmotor functions.[15,21] These projections include prefrontal and posterior parietal cortices, encompassing, at least in human homologue regions, areas associated with language function (Fig. 1A).

Leiner *et al.* proved to be prescient, predicting at the dawn of the neuroimaging era that cerebellar activation would be associated with many cognitive functions.[11] This prediction has been confirmed in countless studies. Controlling for at least overt motor demands, activation in the cerebellum has been reported in a wide range of nonmotor tasks. Indeed, it is harder to find studies that fail to activate the cerebellum than it is to find studies that reveal "cognitive" cerebellar activation, even when the cerebellum is of secondary interest (and, thus, ends up earning just a passing mention in the discussion section). Meta-analyses of the neuroimaging literature reveal consistent cerebellar activation patterns related to working memory, executive function, emotional processing, and across a range of linguistic tasks.[22–26] Of note, these reviews likely underestimate the extent of cerebellar "cognitive" activations since cerebellar coverage is often incomplete in functional imaging studies.

Resting state analyses of functional magnetic resonance imaging (fMRI) data have offered a different approach to identify networks, looking at how blood oxygenation level-dependent (BOLD)-signal fluctuations correlate across the brain in the absence of a task.[27] The picture emerging from functional connectivity studies is generally consistent,[28–35] and a summary picture based on a comprehensive study[28] involving 1000 participants is presented in Figure 1B. Of note for the current review, all of the major cerebral functional networks (e.g., sensorimotor, frontoparietal, and default mode) show connectivity with cerebellar areas. Indeed, the majority of the cerebellum was assigned to cerebral networks associated with cognitive functions, including language.[28]

In summary, converging evidence from evolutionary biology, anatomical studies of cerebrocerebellar connectivity in nonhuman species, and a vast body of human neuroimaging studies all point toward a far broader picture of cerebellar function than the traditional motor account (but see Glickstein,[36,37] for a skeptical view of the "cognitive cerebellum"). However, mechanistic accounts of a cerebellar contribution to cognition, including language, have remained elusive.[38,39] We turn next to a more detailed discussion of cerebellar anatomy and physiology, work that has played a crucial part in the development of theoretical models of cerebellar function.

A computational model based on cerebellar microanatomy

Many of the basic features of cerebellar microanatomy have long been known, with now-classic contributions from Ramon y Cajal,[40] Dow and Moruzzi,[41] and Eccles *et al.*[42] We do not attempt a comprehensive review of cerebellar cell types and circuitry here (see Refs. 43 and 44); rather, we limit our discussion to a brief, and necessarily simplified, overview of the core features of cerebellar inputs and microcircuits (see Fig. 2) that have inspired computational models.

An estimated 40 million mossy fibers[45] constitute one of the two main sources of input to the cerebellum (in comparison, each optic nerve contains about one million). The vast majority of these fibers project from the pontine nuclei of the brainstem, which in turn receive their primary input from the cerebral cortex.[46] After entering the cerebellum, mossy fibers typically split into several collateral branches, innervating spatially distant cerebellar lobules.[46–48] They synapse onto granule cells, the smallest and most numerous type of

neuron in the brain, with cell count estimates ranging from 70 billion[49] to 109 billion.[50] Unlike the ordered functional topography of the cerebral cortex (e.g., the somatosensory and motor homunculi), neurophysiological mapping of somatosensory receptive fields in the cerebellar granule cell layer of rats has revealed a fractured somatotopy, where neighboring patches respond to stimulation of quite different areas of the body.[51,52] Recently, two high-field fMRI studies[53,54] and an electrical mapping study[55] yielded similar results in humans. Indeed, in contrast to the largely segregated sensory processing streams in the cortex,[56] receptive fields in the cerebellum show considerable overlap across different modalities.[54,56,57] Recent studies even find that the input to a single granule cell can be multimodal.[58,59]

The axons of the granule cells, the parallel fibers, synapse on Purkinje cells and interneurons in the cerebellar cortex. The Purkinje cells are the principal neurons of the cerebellum, estimated to number 28 million in adult humans.[50] These large neurons are characterized by their extensive dendritic trees, with each Purkinje cell receiving a vast amount of synaptic input from between 100,000 and 200,000 parallel fibers.[60]

The other main input to the Purkinje cell originates in the inferior olivary nuclei, located in the brainstem. The inferior olive receives excitatory input from several brainstem nuclei, including the red nucleus which receives dense input from the cerebral cortex.[61] It also receives inhibitory input from the cerebellar nuclei (see below). The axons of the inferior olivary nuclei are referred to as climbing fibers, ascending to the cerebellar cortex. In contrast to the convergent pattern of input from the parallel fibers, each Purkinje cell is innervated by a single climbing fiber. However, this climbing fiber makes a vast number of synaptic contacts on the soma and proximal dendrites of Purkinje cells, and has, in fact, been described as the strongest source of synaptic input in the entire nervous system.[43]

Purkinje cells are the sole output neurons of the cerebellar cortex, sending inhibitory projections to a relatively small number of cells in the deep cerebellar nuclei (~600,000).[62, 63] About half of these cells are excitatory and project to the thalamus, the red nucleus and other targets of cerebellar output, with considerable collateralization.[48, 61] The other main group of nuclear cells are inhibitory and project to the inferior olive.[64] In addition to the inhibitory projections from Purkinje cells, the deep cerebellar nuclei also receive excitatory input from mossy fibers and climbing fibers.[43, 65]

In summary, cerebellar microanatomy is characterized by an initial massive divergence of input (~40 million mossy fibers mapping onto ~70–100 billion granule cells), followed by an even more massive convergence (first onto ~28 million Purkinje cells and then onto only ~600.000 cells in the deep nuclei). Notably, and in marked contrast to the ubiquitous reciprocal connections of the cerebral cortex, this information processing takes place in a primarily feed-forward loop[66] involving only a few synapses.

The extensive mapping of cerebellar circuits by Eccles *et al.,*[42] in combination with early models of neural plasticity,[67, 68] inspired computational theories of cerebellar function that remain influential,[69, 70] including the so-called Marr-Albus model. In this model, the divergence from mossy fibers to granule cells served to recode inputs to promote the

discriminability of input patterns. Learning entails a supervised, error-based algorithm in which a teaching signal reduces the synaptic weights on inputs associated with the error. As performance improves, the magnitude of this error signal, and hence, further synaptic weight adjustment, is reduced. Inspired by the physiology and anatomy of the cerebellar cortex, Albus proposed that the climbing fiber outputs serve as an error signal, shaping changes in the parallel fiber– Purkinje cell synapses.[70]

In 1984, Masao Ito extended this idea to provide a more explicit model of how the cerebellum contributed to motor control. Building on classic ideas in engineering,[71] Ito proposed that there were two controls loops (see Fig. 3): An external feedback loop in which the agent interacted with the environment and an internal feed-forward loop that ran through the cerebellum. The internal loop allows the system to operate in a predictive manner. Specifically, processing within this loop provides a mechanism to combine motor commands with information about the current sensory state (or context) to anticipate, or predict the next sensory state. This predictive mechanism has come to be referred to as an internal model, reflecting the idea that the system has learned to simulate the relationship between a motor command and the resulting outcome (for reviews of very similar ideas such as efference copy and corollary discharge, see Refs. 72 and 73). At a more abstract level, internal models can be conceptualized as representations of the context-specific dynamics of an agent's interactions with the environment. In motor control, an internal model might generate predictions of the sensory consequences of an action (e.g., the anticipation of hearing a ringing sound when we finish dialing a phone number); in cognition, an internal model might generate predictions of the consequences of manipulating more abstract representations (e.g., the mental simulation of several possible moves in a game of chess).[13]

An internal model account of classical (Pavlovian) reflex conditioning

How well does the internal model hypothesis match empirical data from neurophysiological experiments? A simple Pavlovian conditioning task, eyeblink conditioning, has served as a model paradigm for studying cerebellar learning.[74] In one version of this preparation, an air puff to the eye serves as the unconditioned stimulus (US), eliciting an instinctive blink (the unconditioned response, UR) to protect the cornea. If this US is repeatedly preceded by a neutral stimulus such as a tone (conditioned stimulus, CS), then the animal will come to blink in anticipation of the US, a conditioned response (CR).[74,75] Moreover, with extended training, the dynamics of the CR will be adjusted such that the blink is maximal around the onset time of the airpuff.[76,77] In the early 1980s, Richard Thompson and colleagues demonstrated that lesions of the cerebellum in the rabbit can selectively abolish the CR with minimal impact on the UR, providing the first demonstration of a localized memory circuit in a mammalian species.[78]

Electrophysiological studies have shown that information about the CS (e.g., the tone) is relayed along the mossy fiber–parallel fiber pathway, generating simple spikes in the Purkinje cell. In contrast, information about the US (e.g., the air puff) is relayed via the climbing fiber pathway and generates complex spikes.[74] Over the course of conditioning, Purkinje cells gradually exhibit adaptively timed pauses in simple spike firing,[74] consistent with the long-term depression (LTD) mechanism proposed by Albus[79] and Ito.[80] Given that

these neurons inhibit the output nuclei of the cerebellum, this pause temporarily releases the nuclei from inhibition, allowing the activation of motor nuclei that control the eye blink.[81]

Thus, we see a predictive system in action: the momentarily disinhibited excitatory output from the cerebellar nuclei signals the anticipation of the aversive stimulus, allowing a precisely timed, predictive eyeblink response. Simultaneously, the nuclei's inhibitory output to the inferior olive ensures that the—now anticipated—aversive stimulus (US) has a decreased probability of eliciting complex spikes.[74] The observation that complex spike probability decreases as the US changes from being unexpected to predicted[82] strongly suggests that they code prediction errors, the primary learning signal posited in the Marr-Albus and internal model hypotheses.

Critically, if the climbing fibers signal a sensory prediction error, nucleoolivary inhibition (signaling a predicted air puff) and the sensory feedback (signaling an actual air puff) must be temporally synchronized. Remarkably, since this is a monosynaptic projection, maximal inhibition of the olive occurs 25–75 ms after cerebellar nuclear cells fire, corresponding to the latency of sensory feedback about the US.[82] Moreover, in order to function as effective teaching signals during conditioning, complex spikes must induce plasticity on parallel fiber– Purkinje cell synapses that *were* active (and failed to predict) at some time in the immediate past. In line with this theoretical requirement, LTD is maximal for parallel fiber–Purkinje cell synapses that were active about 80 ms before the climbing fiber signal.[83]

Reviewing the extensive body of research on eyeblink conditioning in light of the internal model hypothesis, Rasmussen and Hesslow recently concluded that "in an important sense, learning-induced changes in Purkinje cell activity constitute an "expectation" or "anticipation" of a future event (the unconditional stimulus), and, consistent with theoretical models, future learning depends on the accuracy of this expectation" (Ref. 82, p. 116). Thus, at least for classical conditioning of the eyeblink reflex, abstract metaphors such as "predictions," "expectations," and "internal models" map remarkably well onto concrete observable neural activity. We next ask whether this theoretical model can also account for key aspects of the cerebellar role in voluntary motor control and then turn to how these ideas may be relevant for understanding the role of the cerebellum in cognition and language.

Cerebellar internal models in voluntary motor control

Skillful movement relies on the successful resolution of a set of more specific problems, [84] several of which are known to critically involve the cerebellum. First, skilled movement requires precise temporal coordination across different muscles, and a breakdown in this *timing* is a cardinal feature of the motor deficits seen in cerebellar ataxia. These patients are unable to precisely control the timing that is essential for producing rapid movements[85] or coordinate dynamic interactions that arise in multi-joint movements.[86]

Second, relationships between motor commands and their sensory consequences are necessarily variable since both the body and the environment can change over time.[87] For instance, when lifting a glass, the same applied force will produce different sensory consequences depending on whether the glass is full or empty. And over development and

aging, or in the case of disease, the strength of muscles and the dynamics of the body change markedly. Thus, motor behavior must be adaptive to account for these changes. Sensorimotor *adaptation* has been extensively studied in experiments where the visual consequences of a movement are altered: For example, the participant may be asked to make reaching movements while wearing prism glasses[88] or in a virtual environment in which the visual feedback is rotated with respect to the true position of the hand.[89] When encountering such perturbations, participants initially produce large performance errors. These diminish over time as the motor system adapts to this novel environment. A large body of both animal[90] and human[89] research highlights the critical role of the cerebellum in sensorimotor adaptation.

A third fundamental problem in motor control is that, owing to inherent neuronal transmission delays, sensory feedback is too slow to effectively guide movements.[91] In effect, the sensory feedback, once it arrives at the motor control centers, is already out-of-date. Smooth motor control thus relies on a neural mechanism that can predict the sensory consequences of action, allowing motor control to be based on estimates of the *current* state of the organism and its environment. Supporting a role for the cerebellum in sensorimotor *prediction*, patients with cerebellar damage are significantly more impaired on motor tasks requiring predictive control of movement than on tasks relying on feedback control.[92]

In terms of the computational mechanisms involved, sensorimotor adaptation is widely held to rely on the incremental updating of forward (predictive) internal models driven by sensory prediction errors.[84,87] As illustrated in Figure 2, cerebellar internal models can solve the delayed feedback problem by predicting the likely sensory consequences of an executed motor command.[87] Indeed, disabling the predictor module in computer simulations of feedforward control leads to motor control deficits resembling those seen in humans after cerebellar damage[93] and transiently disrupting the cerebellum in healthy individuals with transcranial magnetic stimulation (TMS) results in reaching errors that indicate the movements were based on an estimate of hand position that was about 140 ms out of date.[94] Finally, the temporal requirements for the effective use of internal models in motor control, and the impaired timing seen in patients with cerebellar degeneration, imposes constraints on the *type of predictions* that are likely to involve the cerebellum. Thus, while prediction, in a more general sense, is a key aspect of *all* brain function,[95] the cerebellum appears to be critical for fast online prediction of immediately upcoming events, tuned to the temporal range critical for motor control.[96,97]

In summary, the internal model hypothesis has proven to be a fruitful explanatory construct in addressing several key problems in motor control and has over the last decades become the dominant theoretical model of cerebellar motor function.[98]

Cerebellar internal models beyond motor control: the case of language processing

As reviewed in the introduction, a vast number of imaging studies report cerebellar activations in linguistic (and other cognitive, "beyond-motor") tasks.[22,23,26,99,100] While serious language impairments such as expressive or receptive aphasia are very rare following

acquired cerebellar lesions in adults,[101] a wide range of subtler linguistic deficits have been reported.[38,102] The most prominent of these cerebellar language impairments is motoric in nature, ataxic dysarthria. This disorder is characterized by a marked disruption in the temporal patterns of speech, including slowed rate, decreased range of syllable durations, and increased variability of between-utterance syllable durations.[103,104] Notably, however, deficits in other aspects of language processing are also observed in cerebellar patients. For instance, neuropsychological studies frequently find reduced verbal fluency[101,105,106] and impaired verbal working memory[106–108] following cerebellar lesions. Moreover, there are several reported cases of acquired dyslexia and problems with high-level aspects of language such as grammar/syntax, semantic access, and sentence formation.[38,102,104]

To what extent can a generalization of the internal model hypothesis account for these neuroimaging and clinical findings? In addressing this question, we will organize our discussion around the three themes presented in the previous section concerning how the cerebellum may contribute to motor control. First, cerebellar internal models may contribute to the *timing* of nonmotor as well as motor aspects of language processing.[109] Second, cerebellar internal models may help to continuously *adapt* linguistic (articulatory, auditory, phonological, lexical, syntactic, or semantic) representations to changing internal and environmental contexts. Third, internal models might aid language processing by providing continuous *predictions* of the next linguistic (motor, perceptual, or cognitive/semantic) state.[97] We consider the evidence for each of these properties of internal models in the following sections, examining their relevance for understanding cerebellar involvement in language.

Timing

Language, and in particular speech, is an inherently temporal phenomenon.[109] Indeed, the precise duration and speed of articulatory movements, or pauses, are essential for identifying subtle cues that discriminate between meaning-bearing sounds in many languages.[110] Moreover, a clear temporal structure can also be observed at the level of syllables, which tend to be produced at a rate of 3–6 Hz.[109] Oscillating activity in the auditory cortex entrains to the speech signal in this frequency range, and the degree of entrainment is modulated by speech intelligibility.[111] Extensive neurophysiological work in both animals and humans suggests that this entrainment reflects a mechanism of temporal attention[112–114] in that the maximal excitability of the cortex overlaps with the maximal information content in the acoustic signal.[112]

Inspired by the idea that the cerebellum is recruited in a domain-independent manner to represent precise temporal properties,[96] various groups have looked at cerebellar contributions to timing in speech comprehension.[109,115] Consistent with a generalized cerebellar role in timing, patients with cerebellar degeneration are selectively impaired in distinguishing between phonemes on the basis of temporal cues, while discrimination based on spectral cues is preserved.[115,116] To date, direct evidence for a cerebellar role in temporal attention to language is lacking; however, lesions to the cerebellum reduce predictive sensitivity to the temporal structure of auditory (albeit nonlinguistic) events.[117]

Adaptation

Speech perception must be flexible, allowing us to communicate in a wide range of environments and contexts. While challenging, we are able to talk in noisy environments or understand the heavily accented speech of a non-native speaker.[118] Successful comprehension entails the use of multiple sources of contextual information, including acoustic,[119] visual,[120,121] lexical,[122] and sentential.[123]

Importantly, following repeated exposure to a distorted speech signal in a disambiguating contextual information, we are able to perceive an initially unintelligible signal in the absence of supporting contextual cues.[121,124] And repeated exposure to a speech sound together with a video of an incongruous lip-movement not only shifts the immediate auditory percept toward the visual input (e.g., the McGurk effect[120]), but can also result in lasting auditory recalibration.[125] This flexibility suggests an adaptive process,[126] similar to the spatial recalibration seen in sensorimotor adaptation experiments.[125]

As with the work on timing, surprisingly few studies have examined the role of the cerebellum in language adaptation. One notable exception is an fMRI study by Guediche *et al.*,[126] where activation in several cerebellar regions showed significant changes following perceptual adaptation to acoustically distorted words. Supporting a causal role for the cerebellum, the level of activation in right Crus I correlated with behavioral measures of adaptive plasticity (i.e., recognition of words in severely distorted speech).[118]

Prediction

Speaking is one of the most coordinated, yet effortless actions that humans engage in. The average speaker is capable of transmitting 2–300 words per minute, with the world record exceeding 600. These rapid rates of communication not only require precise coordination for the speaker, but they also challenge the listener who has to decipher complex sounds, identify lexical units that lack a physical boundary, and integrate information over extended phrases to understand the semantic content. Linguistic theorists have long recognized that our ability to communicate in a highly efficient manner comes about because of the redundancy in language: the listener readily anticipates the utterance of the speaker, generating an expectancy of the linguistic message as well as paralinguistic features that might signal the typology of the utterance (e.g., a question) or the pattern (e.g., for turn taking). Indeed, the intervals between turns in a conversation is usually close to 0 ms,[127] strongly suggestive of predictive mechanisms in language processing.

The predictive nature of language has been explored formally in computational models of language production,[128] perception,[129–131] and acquisition.[132–134] This modeling work is paralleled by a large body of empirical studies (for reviews, see Refs. 135 and 136). For instance, it is well established that subjects read predictable words faster than unpredictable words and often skip them altogether during reading.[137] As argued by Pickering and Garrod, people may use internal models in language production and comprehension to predict "what they are about to perceive or to do, in a way that allows them to "get ahead of the game" (Ref. 138, p. 332). In speech production, an internal model may support a comparison between the predicted and actual speech, allowing the output to be adjusted when

discrepancies are detected.[139] Similarly, internal models could facilitate language comprehension through the active prediction of the speaker's next utterance.

Whether the cerebellum contributes to predictive function in the domain of language remains unclear. While the neuropsychological literature does not offer a compelling case for language comprehension impairments in people with cerebellar lesions,[101] it is possible that the deficits are masked by the redundant nature of language. Recently, a number of research groups have examined cerebellar involvement in linguistic predictions, using more analytic experimental tasks and the tools of cognitive neuroscience (for a critical review, see Ref. 140). Argyropoulos and colleagues used TMS to transiently disrupt cerebellar function while healthy young adults performed a range of linguistic tasks.[140–143] In one experiment, cerebellar TMS was found to affect associative linguistic priming where the relationship between the prime and target is based on sequential probabilities (e.g., pigeon-HOLE). Interestingly, there was no effect on semantic priming when the relationship was based on overlapping semantic features (e.g., penny-COIN).

While these results support a cerebellar role in linguistic prediction, they also provide an important constraint in that only predictions involving temporal (or sequential) associations appear to critically rely on the cerebellum. Converging evidence was provided by Lesage *et al.*[144] in an experiment where participants listened to spoken sentences and were required to look, as quickly as possible, at one of four pictures that corresponded to the last word. The sentences either provided a context that strongly predicted the immediately upcoming final word or created a context in which all of the pictures were equally plausible. Crucially, repetitive TMS applied over the right cerebellar hemisphere selectively slowed saccade reaction times in the predictive condition.

Using fMRI, we investigated whether linguistic prediction could provide a functional account of the frequently observed cerebellar activation in linguistic neuroimaging studies, including experiments focusing on higher-level semantic processing.[24,25] In each trial (see Fig. 4A), participants read five sequentially presented words that formed either (1) a coherent sentence with a highly predictable last word (e.g., "two plus two is four"), (2) the beginning of a coherent sentence ending in an unexpected last word (e.g., "[the water] had frozen to cars"), or (3) a random sequence of words (e.g., "fast in clock plane"). Contrasting activity time-locked to the last words in predictable sentences to activity time-locked to the unpredictable last words in the random word sequences revealed a significant cluster in the right cerebellar hemisphere. This pattern can be viewed as reflecting the operation of a linguistic internal model involved in generating a sensory expectancy[145] (see Fig. 4B). In another fMRI study, Lesage *et al.*[146] presented only coherent sentences, while parametrically varying the contextual probability of the final target word (presented at a variable interval after the sentence stem). Consistent with our results, they found that activity in the right posterior cerebellum during the interval between the sentence stem and the target word was positively correlated with the contextual probability of the target word. Thus, while further replications and control experiments are clearly needed,[140] these fMRI results, together with the TMS work presented above, provide promising evidence for a cerebellar role in the generation of linguistic predictions (see Ref. 140 for a discussion of potential methodological limitations and remaining questions, some of which we return to below).

The cerebellum, language, and internal models: taking stock

Although the most prominent language problems associated with cerebellar disorders involve language production,[104,147] the studies reviewed in the previous section also suggest a role for the cerebellum in language comprehension. Building on the sensorimotor control and learning literature, we have highlighted how the concepts of timing, adaptation, and prediction may serve as guides for future theoretical and empirical work on the linguistic cerebellum. While we reviewed each of these concepts separately, they are clearly interlinked. For example, speech adaptation has been explained in terms of a supervised learning mechanism, where the discrepancy between the predicted and the actual auditory input is used to revise an internal model of the context–auditory relationship.[126] Thus, at a general level, we have considered whether the internal model hypothesis offers a useful explanatory construct when addressing the cerebellar role in language processing.

While recent empirical work has provided encouraging support for the extension of the internal model hypothesis from motor control to language processing, it remains to be seen if the similarities are superficial, characteristic of any adaptive system, or arise from specific, computational capabilities of the cerebellum. In considering the latter, we recognize that the extension to the language domain presents some interesting puzzles for current theories of cerebellar function. We review two of these in the final section.

What is the information content of cerebellar linguistic predictions?

One question concerns constraints on predictive functions of the cerebellum in the language domain. Specifically, what kind of information does the cerebellum have access to, and what kind of information can it influence? The TMS and fMRI findings reviewed above are consistent with the hypothesis that the cerebellum is involved in predicting high-level semantic representations. However, the observed (behavioral and hemodynamic) effects could conceivably be explained through a cerebellar influence on more basic sensory or motor functions. For example, semantic information might have been used to predict or prime acoustic or visual representations, articulatory programs, phonological codes or syntax, rather than semantics per se. Indeed, in the internal model account of speech adaptation proposed by Guediche *et al.*,[118] high-level linguistic (lexical) information is used to predict early sensory (prelexical) information. Thus, the existing evidence is compatible with cerebellar predictions at, and between, many levels of language processing, ranging from basic sensory and articulatory processing to semantics.[140] Whether cerebellar predictions are limited to some levels, and whether different linguistic levels show distinct or overlapping cerebellar topographies, are important questions for future research.[140]

Insights into the constraints on cerebellar-based predictions are also likely to benefit from consideration of anatomy, and in particular the projections from the cerebellum to the cerebral cortex. One general conclusion from the nonhuman primate literature is that these appear to be stronger for dorsal processing streams (closely related to motor control) than for ventral processing steams (closely related to object recognition). This pattern of connectivity would be in line with notions of a stronger cerebellar contribution to procedural (articulation, phonetic transitions, syntax) than to declarative (lexical, semantic) aspects of

language processing,[148] and would appear at odds with the idea that the cerebellum generates predictions of high-level semantic content.[143–145]

However, it should be noted that structural[149] and functional[28] connectivity studies in humans suggest more extensive cerebrocerebellar connectivity than has been identified in tracing studies in monkeys[15] (see Fig. 1A and B). It remains unknown if these anatomical discrepancies reflect differences between species or a differential sensitivity/specificity of the respective methods. Importantly, unlike the viral tracing procedure used in the monkey studies, the functional connectivity analyses used in the human neuroimaging studies cannot reliably determine the directionality of cerebrocerebellar connectivity[28] or distinguish between parallel fiber and climbing fiber input to the cerebellar cortex.[150] If all cerebrocerebellar connections were bidirectional, forming closed loops,[15] distinguishing between input and output signals would not be critical. But a close examination of the primate data suggests that this may not always be the case. For example, there is evidence for cerebral projections to the cerebellum from area 7a of the parietal lobe, whereas studies examining cerebellar projections to this same area have yielded negative results.[19] Since the functional interpretation of cerebrocerebellar connectivity critically relies on directional information (recall that *input* to the cerebellum from cognitive areas of the cortex would still be compatible with an exclusive role in motor control as long as cerebellar *output* was solely directed to cerebral motor areas), the interpretation of findings from neuroimaging studies is limited. Consequently, more detailed knowledge of anatomical connectivity in humans (possibly based on novel methods such as polarized light imaging[151,152]) is needed to determine the anatomical plausibility of cerebellar contributions to acoustic, visual, articulatory, phonological, syntactic, and/or semantic processing.

The preceding paragraphs focused on what-predictions, that is, on constraints concerning the representational content of cerebellar processing. However, motivated by the compelling evidence for cerebellar involvement in temporal processing,[96,110] some researchers have argued that cerebellar contributions to language may be best understood in terms of when-predictions, generating expectancies of the temporal properties of information critical for coordinated communication. For example, Kotz and Schwartze[109,153,154] propose that, through learning, the cerebellum can help transform a continuous acoustic input stream into meaningful events, as well as generate and monitor predictions concerning the timing between these events. In this view, deficits related to semantic processing might be secondary to a failure to efficiently prepare extracerebellar regions to anticipate an input at a specific time. A further specification of when-predictions is based on the time scale over which such predictions are made,[154,155] with some imaging evidence suggesting that the right cerebellar hemisphere processes high-pass filtered information (segmental properties important for speech perception), while the left cerebellar hemisphere processes low-pass filtered information (prosodic and melodic properties),[155] thus mirroring the lateralized sensitivity to temporal structure seen in the cerebral cortex.[154]

Importantly, while all predictions generally contain both a what and when component (however rudimentary), the relative importance of representational and temporal fidelity likely varies widely across different neural mechanisms; consider the quick-and-dirty direct amygdala pathway versus the slow-and-accurate cortex–amygdala pathway in threat

perception.[156] Compared to the compelling body of research suggesting a crucial role for the cerebellum in representing when across motor and nonmotor domains,[96,109,110,115,117,153,154,157] the current evidence has little to say in terms of the precision of cerebellar "what" representations. We do note, however, that in all of the semantic priming studies (see section "Cerebellar internal models beyond motor control: the case of language processing" above), the timing of the linguistic stimuli was fixed across experimental conditions, thus making it unlikely that the cerebellar contribution was simply related to predictions of when the target stimulus would be presented. Nonetheless, disentangling the relative importance of what and when for cerebellar-based predictions is clearly an important task for future studies.

A related question concerns the *representational capacity* of cerebellar output. While the input to the cerebellum is certainly massive, it is eventually funneled onto a much more limited number of neurons in the deep nuclei (~600.000). As a point of comparison, the arcuate fasciculus connecting Broca's and Wernicke's areas contains roughly between 15 and 20 million axons (extrapolated from estimates of 0.47 million per mm^2 in the corpus callossum[158]). If this relatively small number of output neurons poses serious limitations on the functional role of the cerebellum in language remains unknown, but should be considered in future theoretical models. For instance, a limited capacity output channel might pose little difficult for the generation of when-predictions (for which fast and reliable signal transmission is more important than bandwidth), but could potentially prohibit more information-dense what-predictions (e.g., detailed semantic content).

What is the nature of the error signal?

Error-based learning has long been a central tenet of computational theories of the cerebellum,[43, 69, 70, 159] and is critical to ideas about how this structure develops and refines internal models.[43, 91, 159] We have drawn on these ideas in considering cerebellar contributions to language. For example, in our previously mentioned fMRI study, we observed widespread activity in the posterior cerebellar hemispheres in response to sentences that terminated with a semantic violation, compared to sentences that were semantically coherent[145] (see Fig. 4C and D). Similarly, the electrophysiological P600 event-related potential, a signature of syntactic violations, was found to be reduced in patients with cerebellar infarctions.[160]

However, this generalization of error-based learning to the domain of language raises an intriguing question concerning the informational content of the error signal. As specified in computational models of sensorimotor adaptation, the error signal, or sensory prediction error "not only tells the system that it missed the goal but also specifies the particular way in which the target was missed" (Ref. 91, p. 742). Information about the direction (and perhaps magnitude) of errors is ideal for a system designed to optimize movements: a reach that results in missing the target to the left should lead to a correction that results in the same context eliciting a movement that is shifted to the right. Gradient descent algorithms ensure that the error signal will be used to improve the output of the system, and should, over time, reach a level of asymptotic performance.[89]

It is hard to imagine corresponding error signals for many aspects of language, at least when considering something like semantics or syntax (or more generally, in most cognitive domains, where the relevant associations may be entirely arbitrary[149,161–163]). Consider the pair of sentences "The water had frozen to ice" and "The water had frozen to cars." While an internal model would surely be useful to detect that "cars" violates a semantic expectancy, it is not clear how the error could be used to refine that internal model, at least not in a way that is analogous to how we think of error signals being used to adapt internal models in sensorimotor control. And if we moved to a world in which water, when cooled, turned into cars rather than ice, there is no continuous space through which this semantic system could gracefully adapt. Instead of a vectorial or directional error signal, the semantic violation appears to be more of a binary signal (i.e., correct or incorrect) or unidimensional signal (i.e., graded only in relation to the degree of "incorrectness," but with no vectorial information about "the particular way in which the target was missed"). Such error signals could still be used to drive associative learning, but they are a very *different kind* of error signal than the one assumed in most models of sensorimotor adaptation.

This point raises an important challenge for models that seek to extend mechanistic ideas from the sensorimotor domain to cognition and will be an important issue to consider in future research. It may be that our quest for a parsimonious account of cerebellar function, one in which similar operations and algorithms are applied to different content domains, is misguided or at least over simplified. Recent reports of considerable physiological heterogeneity across the cerebellar cortex would suggest that the common assumption of a unitary cerebellar function may need to be revised.[164] Perhaps the cerebellum is important for generating predictions for motor control and language, but the manner in which these predictions are shaped may be quite different.

Alternatively, it is important to recognize that the traditional model of cerebellar error signals for motor learning has also been the subject of intense debate. While the standard model holds that errors are encoded in the complex spikes generated by climbing fibers,[165] it has also been noted that the low firing frequency of these neurons (~ 1 Hz) imposes limitations on the utility of these responses for signaling errors.[166] A competing model has consequently emerged, in which directional error signals are hypothesized to be encoded in the high frequency simple spike activity generated by parallel fibers.[166] This high-frequency activity would clearly have the bandwidth necessary to encode detailed information about movements (position, velocity, etc.) and movement errors. It has further been suggested that complex spikes may code the sensitivity to an error rather than directional error information.[167] Such a (nondirectional) error-sensitivity signal would indeed seem more compatible with the kinds of error signals observed in studies of semantic or syntactic violations in language.

This leads to the question of whether sensorimotor adaptation is driven by nondirectional error signals? Perhaps. In one scenario, the output (or predictions) of any given internal model is fixed with respect to its information content (or the events they code) and can only be modified in a binary manner (i.e., more or less probable, or more or less intense). However, the vast representational capacity of the granule cell layer might allow very similar contextual representations to connect to very different output representations in the nuclei.

Thus, adaptation would be a process in which novel context–output mappings are learned, while old ones are unlearned, or extinguished. Interestingly, the characteristic cerebellar fractured somatotopy of the cerebellum,[51],[53–55] in which neighboring cortical patches represent distal body parts or even different sensory modalities, may be optimal for the rapid formation of "new associations between movements and relatively arbitrary, remote sensory consequences."

After a perturbation of sensory feedback (e.g., through wearing prism glasses), some events or sensory consequences would become less likely given a certain context, while other events (consequently involving *different* internal models) would become more likely given this same context. Importantly, while each internal model could be updated using binary error signals related to the accuracy of a specific predicted sensory consequence, gradual changes in the summed or population activity across many internal models may still take on the directional properties assumed in computational models of error-based sensorimotor adaptation.

Conclusions

An increasingly influential theory in the cognitive neuroscience of motor control holds that the cerebellum implements *internal models*, neural representations of the context-specific dynamic properties of agent/environment interactions. In addition to being grounded in cerebellar microanatomy and physiology, this hypothesis has provided a parsimonious account of several key aspects of motor control. Thus, the search for a unified account of cerebellar function across both motor and non-motor domains has led to efforts to generalize this computational idea to language processing. While initial experimental results are promising, a number of key questions remain, concerning both the information content of cerebellar predictions and the nature of error signals. Irrespective of whether our speculative solution to the latter problem will prove to have merit, we hope to have shown that results from studies of the "cognitive cerebellum" may well feedback on, and indeed challenge, established models of the cerebellar role in motor control.

Acknowledgments

References

1. Rolando, L. Saggio sopra la vera struttura del cervello dell'uomo e degli animali e sopra le funzioni del sistema nervoso. Sassari: Stampeia da S.S.R.M. Privilegiata; 1809.

2. Holmes G. The cerebellum of man. Brain. 1939; 62:1–30.

3. Penfield W, Perot P. The brain's record of auditory and visual experience. A final summary and discussion. Brain. 1963; 86:595–696. [PubMed: 14090522]

4. Rangarajan V, et al. Electrical stimulation of the left and right human fusiform gyrus causes different effects in conscious face perception. J. Neurosci. 2014; 34:12828–12836. [PubMed: 25232118]

5. Salminen-Vaparanta N, et al. Subjective characteristics of TMS-induced phosphenes originating in human V1 and V2. Cereb. Cortex. 2014; 24:2751–2760. [PubMed: 23696280]

6. Riklan M, Cullinan T, Shulman M, Cooper IS. A psychometric study of chronic cerebellar stimulation in man. Biol. Psychiatry. 1976; 11:543–574. [PubMed: 786383]

7. Tomlinson SP, Davis NJ, Bracewell RM. Brain stimulation studies of non-motor cerebellar function: a systematic review. Neurosci. Biobehav. Rev. 2013; 37:766–789. [PubMed: 23500608]

8. Snider RS, Wetzel N. Electroencephalographic changes induced by stimulation of the cerebellum of man. Electroencephalogr. Clin. Neurophysiol. 1965; 18:176–183. [PubMed: 14255044]

9. Galliano E, De Zeeuw CI. Questioning the cerebellar doctrine. Prog. Brain Res. 2014; 210:59–77. [PubMed: 24916289]

10. Parvizi J. Corticocentric myopia: old bias in new cognitive sciences. Trends Cogn. Sci. 2009; 13:354–359. [PubMed: 19595625]

11. Leiner HC, Leiner AL, Dow RS. Does the cerebellum contribute to mental skills? Behav. Neurosci. 1986; 100:443–454. [PubMed: 3741598]

12. Ramnani N. The primate cortico-cerebellar system: anatomy and function. Nat. Rev. Neurosci. 2006; 7:511–522. [PubMed: 16791141]

13. Ito M. Control of mental activities by internal models in the cerebellum. Nat. Rev. Neurosci. 2008; 9:304–313. [PubMed: 18319727]

14. Buckner RL. The cerebellum and cognitive function: 25 years of insight from anatomy and neuroimaging. Neuron. 2013; 80:807–815. [PubMed: 24183029]

15. Strick PL, Dum RP, Fiez JA. Cerebellum and nonmotor function. Annu. Rev. Neurosci. 2009; 32:413–434. [PubMed: 19555291]

16. Leiner HC, Leiner AL, Dow RS. Cognitive and language functions of the human cerebellum. Trends Neurosci. 1993; 16:444–447. [PubMed: 7507614]

17. Balsters JH, et al. Evolution of the cerebellar cortex: the selective expansion of prefrontal-projecting cerebellar lobules. Neuroimage. 2010; 49:2045–2052. [PubMed: 19857577]

18. Barton RA, Venditti C. Rapid evolution of the cerebellum in humans and other great apes. Curr. Biol. 2014; 24:2440–2444. [PubMed: 25283776]

19. Bostan AC, Dum RP, Strick PL. Cerebellar networks with the cerebral cortex and basal ganglia. Trends Cogn. Sci. 2013; 17:241–254. [PubMed: 23579055]

20. Glickstein M, May JG, Mercier BE. Corticopontine projection in the macaque: the distribution of labelled cortical cells after large injections of horseradish peroxidase in the pontine nuclei. J. Comp. Neurol. 1985; 235:343–359. [PubMed: 3998215]

21. Bostan, AC., Strick, PL. Cerebellar outputs in non-human primates: an anatomical perspective using transsynaptic tracers. In: Manto, M.Gruol, DL.Schmahmann, J., et al., editors. Handbook of the Cerebellum and Cerebellar Disorders. Dordrecht: Springer; 2013. p. 549-569.

22. Keren-Happuch E, Chen S-HA, Ho M-HR, Desmond JE. A meta-analysis of cerebellar contributions to higher cognition from PET and fMRI studies. Hum. Brain Mapp. 2014; 35:593–615. [PubMed: 23125108]

23. Fedorenko E, Hsieh P-J, Nieto-Castannón A, et al. New method for fMRI investigations of language: defining ROIs functionally in individual subjects. J. Neurophysiol. 2010; 104:1177–1194. [PubMed: 20410363]

24. Fedorenko E, Behr MK, Kanwisher N. Functional specificity for high-level linguistic processing in the human brain. Proc. Natl. Acad. Sci. U.S.A. 2011; 108:16428–16433. [PubMed: 21885736]

25. Price CJ. A review and synthesis of the first 20years of PET and fMRI studies of heard speech, spoken language and reading. Neuroimage. 2012; 62:816–847. [PubMed: 22584224]

26. Stoodley CJ, Schmahmann JD. Functional topography in the human cerebellum: a meta-analysis of neuroimaging studies. Neuroimage. 2009; 44:489–501. [PubMed: 18835452]

27. Biswal B, Yetkin FZ, Haughton VM, Hyde JS. Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. Magn. Reson. Med. 1995; 34:537–541. [PubMed: 8524021]

28. Buckner RL, Krienen FM, Castellanos A, et al. The organization of the human cerebellum estimated by intrinsic functional connectivity. J. Neurophysiol. 2011; 106:2322–2345. [PubMed: 21795627]

29. Sang L, et al. Resting-state functional connectivity of the vermal and hemispheric subregions of the cerebellum with both the cerebral cortical networks and subcortical structures. Neuroimage. 2012; 61:1213–1225. [PubMed: 22525876]

30. Habas C, et al. Distinct cerebellar contributions to intrinsic connectivity networks. J. Neurosci. 2009; 29:8586–8594. [PubMed: 19571149]

31. O'Reilly JX, Beckmann CF, Tomassini V, et al. Distinct and overlapping functional zones in the cerebellum defined by resting state functional connectivity. Cereb. Cortex. 2010; 20:953–965. [PubMed: 19684249]

32. Krienen FM, Buckner RL. Segregated frontocerebellar circuits revealed by intrinsic functional connectivity. Cereb. Cortex. 2009; 19:2485–2497. [PubMed: 19592571]

33. Bernard JA, et al. Resting state cortico-cerebellar functional connectivity networks: a comparison of anatomical and self-organizing map approaches. Front. Neuroanat. 2012; 6:31. [PubMed: 22907994]

34. Bernard JA, et al. Dissociable functional networks of the human dentate nucleus. Cereb. Cortex. 2014; 24:2151–2159. [PubMed: 23513045]

35. Kipping JA, et al. Overlapping and parallel cerebello-cerebral networks contributing to sensorimotor control: an intrinsic functional connectivity study. Neuroimage. 2013; 83:837–848. [PubMed: 23872155]

36. Glickstein M. What does the cerebellum really do? Curr. Biol. 2007; 17:R824–R827. [PubMed: 17925205]

37. Glickstein M, Doron K. Cerebellum: connections and functions. Cerebellum. 2008; 7:589–594. [PubMed: 19002543]

38. Mariën P, et al. Consensus paper: language and the cerebellum: an ongoing enigma. Cerebellum. 2014; 13:386–410. [PubMed: 24318484]

39. De Smet HJ, Paquier P, Verhoeven J, Mariën P. The cerebellum: its role in language and related cognitive and affective functions. Brain Lang. 2013; 127:334–342. [PubMed: 23333152]

40. Ramon y Cajal, S. Histologie du Système Nerveux de l'Homme et des Vertèebrè s. Paris: Maloine; 1911.

41. Dow, RS., Moruzzi, G. The Physiology and Pathology of the Cerebellum. Minneapolis: The University of Minnesota Press; 1958.

42. Eccles, JC., Ito, M., Szentagothai, J. The Cerebellum as a Neuronal Machine. New York: Springer; 1967.

43. Ito, M. Cerebellum. Upper Saddle River, NJ: FT Press; 2012.

44. Manto, M., Gruol, D., Schmahmann, J., et al. Handbook of the Cerebellum and Cerebellar Disorders. Springer; 2012.

45. Tomasch J. The numerical capacity of the human cortico-pontocerebellar system. Brain Res. 1969; 13:476–484. [PubMed: 5305681]

46. Brodal P, Bjaalie JG. Salient anatomic features of the cortico-pontocerebellar pathway. Prog. Brain Res. 1997; 114:227–249. [PubMed: 9193147]

47. Voogd J, Pardoe J, Ruigrok TJH, Apps R. The distribution of climbing and mossy fiber collateral branches from the copula pyramidis and the paramedian lobule: congruence of climbing fiber cortical zones and the pattern of zebrin banding within the rat cerebellum. J. Neurosci. 2003; 23:4645–4656. [PubMed: 12805304]

48. Ruigrok TJH. Ins and outs of cerebellar modules. Cerebellum. 2011; 10:464–474. [PubMed: 20232190]

49. Lange W. Cell number and cell density in the cerebellar cortex of man and some other mammals. Cell Tissue Res. 1975; 157:115–124. [PubMed: 804353]

50. Andersen BB, Gundersen HJG, Pakkenberg B. Aging of the human cerebellum: a stereological study. J. Comp. Neurol. 2003; 466:356–365. [PubMed: 14556293]

51. Shambes GM, Gibson JM, Welker W. Fractured somatotopy in granule cell tactile areas of rat cerebellar hemispheres revealed by micromapping. Brain Behav. Evol. 1978; 15:94–140. [PubMed: 638731]

52. Bower JM. Functional implications of tactile projection patterns to the lateral hemispheres of the cerebellum of the albino rat: the legacy of Wally Welker. Ann. N.Y. Acad. Sci. 2011; 1225:130–141. [PubMed: 21535000]

53. van der Zwaag W, et al. Digit somatotopy in the human cerebellum: a 7T fMRI study. Neuroimage. 2013; 67:354–362. [PubMed: 23238433]

54. Wiestler T, McGonigle DJ, Diedrichsen J. Integration of sensory and motor representations of single fingers in the human cerebellum. J. Neurophysiol. 2011; 105:3042–3053. [PubMed: 21471398]

55. Mottolese C, et al. Mapping motor representations in the human cerebellum. Brain. 2013; 136:330–342. [PubMed: 22945964]

56. Proville RD, et al. Cerebellum involvement in cortical sensorimotor circuits for the control of voluntary movements. Nat. Neurosci. 2014; 17:1233–1239. [PubMed: 25064850]

57. Snider RS, Stowell A. Receiving areas of the tactile, auditory and visual systems in the cerebellum. J. Neurophysiol. 1944; 7:331–357.

58. Sawtell NB. Multimodal integration in granule cells as a basis for associative plasticity and sensory prediction in a cerebellum-like circuit. Neuron. 2010; 66:573–584. [PubMed: 20510861]

59. Huang C-C, et al. Convergence of pontine and proprioceptive streams onto multimodal cerebellar granule cells. Elife. 2013; 2:e00400. [PubMed: 23467508]

60. Harvey RJ, Napper RM. Quantitative studies on the mammalian cerebellum. Prog. Neurobiol. 1991; 36:437–463. [PubMed: 1947173]

61. Voogd, J., Shinoda, Y., Ruigrok, TJH., Sugihara, I. Cerebellar nuclei and the inferior olivary nuclei: organization and connections. In: Manto, M.Gruol, DL.Schmahmann, J., et al., editors. Handbook of the Cerebellum and Cerebellar Disorders. Dordrecht: Springer; 2013. p. 377-436.

62. Heidary H, Tomasch J. Neuron numbers and perikaryon areas in the human cerebellar nuclei. Acta Anat. (Basel). 1969; 74:290–296. [PubMed: 5383711]

63. Hayaran A, Wadhwa S, Gopinath G, Bijlani V. Developing dentate nucleus in man: a qualitative and quantitative study. Exp. Brain Res. 1992; 89:640–648. [PubMed: 1644126]

64. Bengtsson F, Hesslow G. Cerebellar control of the inferior olive. Cerebellum. 2006; 5:7–14. [PubMed: 16527758]

65. Bengtsson F, Ekerot C-F, Jörntell H. *In vivo* analysis of inhibitory synaptic inputs and rebounds in deep cerebellar nuclear neurons. PLoS One. 2011; 6:e18822. [PubMed: 21552556]

66. Buzsaki, G. Rhythms of the Brain. New York: Oxford University Press; 2006.

67. Hebb, DO. The Organization of Behavior. New York: Wiley; 1949.

68. Rosenblatt F. The perceptron: a probabilistic model for information storage and organization in the brain. Psychol. Rev. 1958; 65:386–408. [PubMed: 13602029]

69. Marr D. A theory of cerebellar cortex. J. Physiol. 1969; 202:437–470. [PubMed: 5784296]

70. Albus JS. A theory of cerebellar function. Math. Biosci. 1971; 10:25–61.

71. Bennett S. A brief history of automatic control. IEEE Control Systems Magazine. 1996

72. Grüsser, OJ. On the history of the ideas of efference copy and reafference. In: Debru, C., editor. Essays in the history of the physiological sciences. Proceedings of a Network Symposium of the European Association for the History of Medicine and Health, held at the University Louis Pasteur, Strasbourg, on March 26–27, 1993. Amsterdam and Atlanta, GA: Rodopi B.V.; 1995. p. 35-56.

73. Crapse TB, Sommer MA. Corollary discharge across the animal kingdom. Nat. Rev. Neurosci. 2008; 9:587–600. [PubMed: 18641666]

74. Rasmussen A, Jirenhed D-A, Wetmore DZ, Hesslow G. Changes in complex spike activity during classical conditioning. Front. Neural Circuits. 2014; 8:90. [PubMed: 25140129]

75. Pavlov, IP. Conditioned Reflexes: An Investigation of the Physiological Activity of the Cerebral Cortex. Oxford: Oxford University Press; 1927.

76. Ebel HC, Prokasy WF. Classical eyelid conditioning as a function of sustained and shifted interstimulus intervals. J. Exp. Psychol. 1963; 65:52.

77. Boneau CA. The interstimulus interval and the latency of the conditioned eyelid response. J. Exp. Psychol. 1958; 56:464–471. [PubMed: 13611170]

78. McCormick DA, et al. The engram found? Role of the cerebellum in classical conditioning of nictitating membrane and eyelid responses. Bull. Psychon. Soc. 1981; 18:103–105.

79. Albus JS, Branch DT. A theory of cerebellar function. Math. Biosci. 1971; 10:25–61.

80. Ito, M. The Cerebellum and Neural Control. New York: Raven Press; 1984.

81. Jirenhed D-A, Bengtsson F, Hesslow G. Acquisition, extinction, and reacquisition of a cerebellar cortical memory trace. J. Neurosci. 2007; 27:2493–2502. [PubMed: 17344387]

82. Rasmussen A, Hesslow G. Feedback control of learning by the cerebello-olivary pathway. Prog. Brain Res. 2014; 210:103–119. [PubMed: 24916291]

83. Safo P, Regehr WG. Timing dependence of the induction of cerebellar LTD. Neuropharmacology. 2008; 54:213–218. [PubMed: 17669443]

84. Wolpert DM, Ghahramani Z. Computational principles of movement neuroscience. Nat. Neurosci. 2000; 3:1212–1217. [PubMed: 11127840]

85. Berardelli A, et al. Single-joint rapid arm movements in normal subjects and in patients with motor disorders. Brain. 1996; 119(Pt. 2):661–674. [PubMed: 8800955]

86. Bastian AJ, Martin TA, Keating JG, Thach WT. Cerebellar ataxia: abnormal control of interaction torques across multiple joints. J. Neurophysiol. 1996; 76:492–509. [PubMed: 8836239]

87. Shadmehr R, Smith MA, Krakauer JW. Error correction, sensory prediction, and adaptation in motor control. Annu. Rev. Neurosci. 2010; 33:89–108. [PubMed: 20367317]

88. Martin TA, Keating JG, Goodkin HP, et al. Throwing while looking through prisms. I. Focal olivocerebellar lesions impair adaptation. Brain. 1996; 119(Pt. 4):1183–1198. [PubMed: 8813282]

89. Taylor JA, Ivry RB. Cerebellar and prefrontal cortex contributions to adaptation, strategies, and reinforcement learning. Prog. Brain Res. 2014; 210:217–253. [PubMed: 24916295]

90. Dash S, Thier P. Cerebellum-dependent motor learning: lessons from adaptation of eye movements in primates. Prog. Brain Res. 2014; 210:121–155. [PubMed: 24916292]

91. Wolpert DM, Diedrichsen J, Flanagan JR. Principles of sensorimotor learning. Nat. Rev. Neurosci. 2011; 12:739–751. [PubMed: 22033537]

92. Bastian AJ. Learning to predict the future: the cerebellum adapts feedforward movement control. Curr. Opin. Neurobiol. 2006; 16:645–649. [PubMed: 17071073]

93. Miall, RC., Wolpert, DM. "The cerebellum as a predictive model of the motor system: a Smith predictor hypothesis. In: Ferrell, WR., Proske, U., editors. Neural Control of Movements. New York: Plenum Press; 1995. p. 215-223.

94. Miall RC, Christensen LOD, Cain O, Stanley J. Disruption of state estimation in the human lateral cerebellum. PLoS Biol. 2007; 5:e316. [PubMed: 18044990]

95. Friston K. The free-energy principle: a unified brain theory? Nat. Rev. Neurosci. 2010; 11:127–138. [PubMed: 20068583]

96. Ivry RB, Spencer RMC. The neural representation of time. Curr. Opin. Neurobiol. 2004; 14:225–232. [PubMed: 15082329]

97. Ghajar J, Ivry RB. The predictive brain state: asynchrony in disorders of attention? Neuroscientist. 2009; 15:232–242. [PubMed: 19074688]

98. Ebner, TJ. Cerebellum and internal models. In: Manto, M.Gruol, DL.Schmahmann, J., et al., editors. Handbook of the Cerebellum and Cerebellar Disorders. Dordrecht: Springer; 2013. p. 1279-1295.

99. Fedorenko E, Nieto-Castañnón A, Kanwisher N. Lexical and syntactic representations in the brain: an fMRI investigation with multi-voxel pattern analyses. Neuropsychologia. 2012; 50:499–513. [PubMed: 21945850]

100. Fedorenko E, Nieto-Castañnón A, Kanwisher N. Syntactic processing in the human brain: what we know, what we don't know, and a suggestion for how to proceed. Brain Lang. 2012; 120:187–207. [PubMed: 21334056]

101. Alexander MP, Gillingham S, Schweizer T, Stuss DT. Cognitive impairments due to focal cerebellar injuries in adults. Cortex. 2012; 48:980–990. [PubMed: 21549360]

102. Murdoch BE. The cerebellum and language: historical perspective and review. Cortex. 2010; 46:858–868. [PubMed: 19828143]

103. Ackermann H, Hertrich I. Speech rate and rhythm in cerebellar dysarthria: an acoustic analysis of syllabic timing. Folia Phoniatr. Logop. 1994; 46:70–78. [PubMed: 8173615]

104. Bodranghien F, et al. Consensus paper: revisiting the symptoms and signs of cerebellar syndrome. Cerebellum. 2016; 15:369–391. [PubMed: 26105056]

105. Schweizer TA, Alexander MP, Susan Gillingham BA, et al. Lateralized cerebellar contributions to word generation: a phonemic and semantic fluency study. Behav. Neurol. 2010; 23:31–37. [PubMed: 20714059]

106. Peterburs J, Bellebaum C, Koch B, et al. Working memory and verbal fluency deficits following cerebellar lesions: relation to interindividual differences in patient variables. Cerebellum. 2010; 9:375–383. [PubMed: 20387024]

107. Kirschen MP, et al. Verbal memory impairments in children after cerebellar tumor resection. Behav. Neurol. 2008; 20:39–53. [PubMed: 19491473]

108. Ravizza SM, et al. Cerebellar damage produces selective deficits in verbal working memory. Brain. 2006; 129:306–320. [PubMed: 16317024]

109. Schwartze M, Kotz SA. Contributions of cerebellar event-based temporal processing and preparatory function to speech perception. Brain Lang. 2015 pii: S0093-934X(15)00182-0.

110. Ivry RB, Spencer RM, Zelaznik HN, et al. The cerebellum and event timing. Ann. N.Y. Acad. Sci. 2002; 978:302–317. [PubMed: 12582062]

111. Peelle JE, Gross J, Davis MH. Phase-locked responses to speech in human auditory cortex are enhanced during comprehension. Cereb. Cortex. 2013; 23:1378–1387. [PubMed: 22610394]

112. Morillon B, Hackett TA, Kajikawa Y, Schroeder CE. Predictive motor control of sensory dynamics in auditory active sensing. Curr. Opin. Neurobiol. 2015; 31:230–238. [PubMed: 25594376]

113. Lakatos P, et al. The spectrotemporal filter mechanism of auditory selective attention. Neuron. 2013; 77:750–761. [PubMed: 23439126]

114. Besle J, et al. Tuning of the human neocortex to the temporal dynamics of attended events. J. Neurosci. 2011; 31:3176–3185. [PubMed: 21368029]

115. Ackermann H, Gräber S, Hertrich I, Daum I. Categorical speech perception in cerebellar disorders. Brain Lang. 1997; 60:323–331. [PubMed: 9344481]

116. Ackermann H, Mathiak K, Ivry RB. Temporal organization of 'internal speech' as a basis for cerebellar modulation of cognitive functions. Behav. Cogn. Neurosci. Rev. 2004; 3:14–22. [PubMed: 15191639]

117. Kotz SA, Stockert A, Schwartze M. Cerebellum, temporal predictability and the updating of a mental model. Philos. Trans. R. Soc. Lond. B Biol. Sci. 2014; 369:20130403. [PubMed: 25385781]

118. Guediche S, Holt LL, Laurent P, et al. Evidence for cerebellar contributions to adaptive plasticity in speech perception. Cereb. Cortex. 2015; 25:1867–1877. [PubMed: 24451660]

119. Holt LL. Temporally nonadjacent nonlinguistic sounds affect speech categorization. Psychol. Sci. 2005; 16:305–312. [PubMed: 15828978]

120. McGurk H, MacDonald J. Hearing lips and seeing voices. Nature. 1976; 264:746–748. [PubMed: 1012311]

121. Vroomen J, van Linden S, de Gelder B, Bertelson P. Visual recalibration and selective adaptation in auditory–visual speech perception: contrasting build-up courses. Neuropsychologia. 2007; 45:572–577. [PubMed: 16530233]

122. Ganong WF. Phonetic categorization in auditory word perception. J. Exp. Psychol. Hum. Percept. Perform. 1980; 6:110–125. [PubMed: 6444985]

123. Ladefoged P. Information conveyed by vowels. J. Acoust. Soc. Am. 1957; 29:98.

124. Norris D, McQueen JM, Cutler A. Perceptual learning in speech. Cogn. Psychol. 2003; 47:204–238. [PubMed: 12948518]

125. Bertelson P, Vroomen J, de Gelder B. Visual recalibration of auditory speech identification: a McGurk aftereffect. Psychol. Sci. 2003; 14:592–597. [PubMed: 14629691]

126. Guediche S, Blumstein SE, Fiez JA, Holt LL. Speech perception under adverse conditions: insights from behavioral, computational, and neuroscience research. Front. Syst. Neurosci. 2014; 7:126. [PubMed: 24427119]

127. Wilson M, Wilson TP. An oscillator model of the timing of turn-taking. Psychon. Bull. Rev. 2005; 12:957–968. [PubMed: 16615316]

128. Hickok G. Computational neuroanatomy of speech production. Nat. Rev. Neurosci. 2012; 13:135–145. [PubMed: 22218206]

129. Dikker S, Pylkkänen L. Predicting language: MEG evidence for lexical preactivation. Brain Lang. 2013; 127:55–64. [PubMed: 23040469]

130. Conway CM, Bauernschmidt A, Huang SS, Pisoni DB. Implicit statistical learning in language processing: word predictability is the key. Cognition. 2010; 114:356–371. [PubMed: 19922909]

131. Tian X, Poeppel D. Mental imagery of speech: linking motor and perceptual systems through internal simulation and estimation. Front. Hum. Neurosci. 2012; 6:314. [PubMed: 23226121]

132. Romberg AR, Saffran JR. Statistical learning and language acquisition. Wiley Interdiscip. Rev. Cogn. Sci. 2010; 1:906–914. [PubMed: 21666883]

133. Romberg AR, Saffran JR. Expectancy learning from probabilistic input by infants. Front. Psychol. 2012; 3:610. [PubMed: 23439947]

134. Romberg AR, Saffran JR. All together now: concurrent learning of multiple structures in an artificial language. Cogn. Sci. 2013; 37:1290–1320. [PubMed: 23772795]

135. Pickering MJ, Garrod S. Do people use language production to make predictions during comprehension? Trend Cogn. Sci. 2007; 11:105–110.

136. Kutas M, Federmeier KD. Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). Annu. Rev. Psychol. 2011; 62:621–647. [PubMed: 20809790]

137. Ehrlich SF, Rayner K. Contextual effects on word perception and eye movements during reading. J. Verbal Learn. Verbal Behav. 1981; 20:641–655.

138. Pickering MJ, Garrod S. An integrated theory of language production and comprehension. Behav. Brain Sci. 2013; 36:329–347. [PubMed: 23789620]

139. Tourville JA, Reilly KJ, Guenther FH. Neural mechanisms underlying auditory feedback control of speech. Neuroimage. 2008; 39:1429–1443. [PubMed: 18035557]

140. Argyropoulos GPD. The cerebellum, internal models and prediction in 'non-motor' aspects of language: a critical review. Brain Lang. 2015 pii: S0093-934X(15)00174-1.

141. Argyropoulos GP. Cerebellar theta-burst stimulation selectively enhances lexical associative priming. Cerebellum. 2011; 10:540–550. [PubMed: 21451999]

142. Argyropoulos GP, Kimiskidis VK, Papagiannopoulos S. θ-Burst stimulation of the right neocerebellar vermis selectively disrupts the practice-induced acceleration of lexical decisions. Behav. Neurosci. 2011; 125:724–734. [PubMed: 21942434]

143. Argyropoulos GP, Muggleton NG. Effects of cerebellar stimulation on processing semantic associations. Cerebellum. 2013; 12:83–96. [PubMed: 22752996]

144. Lesage E, Morgan BE, Olson AC, et al. Cerebellar rTMS disrupts predictive language processing. Curr. Biol. 2012; 22:R794–R795. [PubMed: 23017990]

145. Moberget T, Gullesen EH, Andersson S, et al. Generalized role for the cerebellum in encoding internal models: evidence from semantic processing. J. Neurosci. 2014; 34:2871–2878. [PubMed: 24553928]

146. Lesage, E., Hansen, PC., Miall, RC. Cerebellar BOLD response to linguistic stimuli is modulated by predictability. Program No. 733.04. 2014 Neuroscience Meeting Planner. Washington, DC: Society for Neuroscience; 2014. Online

147. Spencer KA, Slocomb DL. The neural basis of ataxic dysarthria. Cerebellum. 2007; 6:58–65. [PubMed: 17366266]

148. Ullman MT. A neurocognitive perspective on language: the declarative/procedural model. Nat. Rev. Neurosci. 2001; 2:717–726. [PubMed: 11584309]

149. Ramnani N, et al. The evolution of prefrontal inputs to the cortico-pontine system: diffusion imaging evidence from Macaque monkeys and humans. Cereb. Cortex. 2006; 16:811–818. [PubMed: 16120793]

150. Diedrichsen J, Verstynen T, Schlerf J, Wiestler T. Advances in functional imaging of the human cerebellum. Curr. Opin. Neurol. 2010; 23:382–387. [PubMed: 20581682]

151. Caspers S, et al. Target sites for transcallosal fibers in human visual cortex—a combined diffusion and polarized light imaging study. Cortex. 2015; 72:40–53. [PubMed: 25697048]

152. Axer M, et al. A novel approach to the human connectome: ultra-high resolution mapping of fiber tracts in the brain. Neuroimage. 2011; 54:1091–1101. [PubMed: 20832489]

153. Schwartze M, Kotz SA. A dual-pathway neural architecture for specific temporal prediction. Neurosci. Biobehav. Rev. 2013; 37:2587–2596. [PubMed: 23994272]

154. Kotz SA, Schwartze M. Cortical speech processing unplugged: a timely subcortico-cortical framework. Trends Cogn. Sci. 2010; 14:392–399. [PubMed: 20655802]

155. Callan DE, Kawato M, Parsons L, Turner R. Speech and song: the role of the cerebellum. Cerebellum. 2007; 6:321–327. [PubMed: 17853077]

156. Rogan MT, LeDoux JE. Emotion: systems, cells, synaptic plasticity. Cell. 1996; 85:469–475. [PubMed: 8653782]

157. Ackermann H. Cerebellar contributions to speech production and speech perception: psycholinguistic and neurobiological perspectives. Trends Neurosci. 2008; 31:265–272. [PubMed: 18471906]

158. Aboitiz F, Scheibel AB, Fisher RS, Zaidel E. Fiber composition of the human corpus callosum. Brain Res. 1992; 598:143–153. [PubMed: 1486477]

159. Doya K. What are the computations of the cerebellum, the basal ganglia and the cerebral cortex? Neural Netw. 1999; 12:961–974. [PubMed: 12662639]

160. Adamaszek M, Strecker K, Kessler C. Impact of cerebellar lesion on syntactic processing evidenced by event-related potentials. Neurosci. Lett. 2012; 512:78–82. [PubMed: 22343021]

161. Ramnani N. Automatic and controlled processing in the corticocerebellar system. Prog. Brain Res. 2014; 210:255–285. [PubMed: 24916296]

162. Balsters JH, Ramnani N. Cerebellar plasticity and the automation of first-order rules. J. Neurosci. 2011; 31:2305–2312. [PubMed: 21307266]

163. Balsters JH, Whelan CD, Robertson IH, Ramnani N. Cerebellum and cognition: evidence for the encoding of higher order rules. Cereb. Cortex. 2013; 23:1433–1443. [PubMed: 22617850]

164. Witter L, De Zeeuw CI. Regional functionality of the cerebellum. Curr. Opin. Neurobiol. 2015; 33:150–155. [PubMed: 25884963]

165. Ito M. Error detection and representation in the olivo-cerebellar system. Front. Neural Circuits. 2013; 7:1. [PubMed: 23440175]

166. Popa LS, Streng ML, Hewitt AL, Ebner TJ. The errors of our ways: understanding error representations in cerebellar-dependent motor learning. Cerebellum. 2016; 15:93–103. [PubMed: 26112422]

167. Marko MK, Haith AM, Harran MD, Shadmehr R. Sensitivity to prediction error in reach adaptation. J. Neurophysiol. 2012; 108:1752–1763. [PubMed: 22773782]

168. Yeo BT, et al. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. J. Neurophysiol. 2011; 106:1125–1165. [PubMed: 21653723]

169. D'Angelo E, Casali S. Seeking a unified framework for cerebellar function and dysfunction: from circuit operations to cognition. Front. Neural Circuits. 2013; 6:116. [PubMed: 23335884]
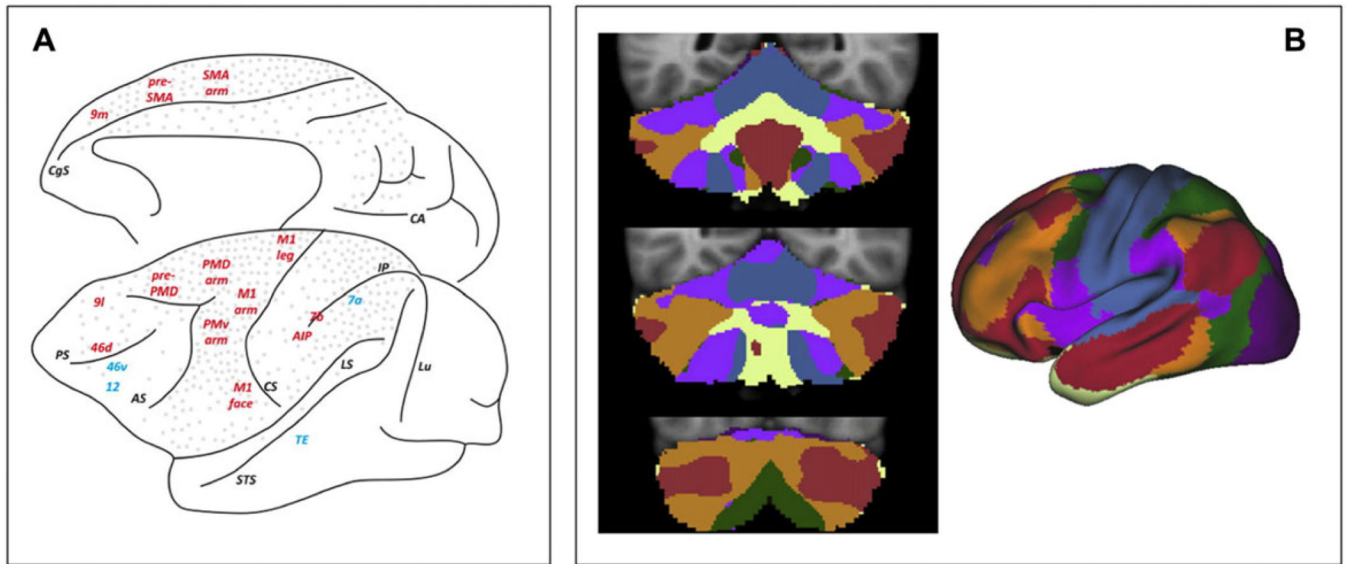
Figure 1.
(A) Cerebrocerebellar connectivity mapped in the monkey with transneuronal viral tracers. Gray dots mark cerebral areas projecting to the cerebellum, and areas printed in red receive projections from the cerebellum. Blue indicates areas that appear to not receive projections from the cerebellum. Adapted, with permission, from Ref. 19. (B) Human cerebrocerebellar connectivity mapped with fMRI. A parcellation of the cerebral cortex into seven large-scale functional networks[168] is shown on the right, while cerebellar areas functionally connected to each cerebral network are displayed on the left. Reproduced, with permission, from Ref. 14.
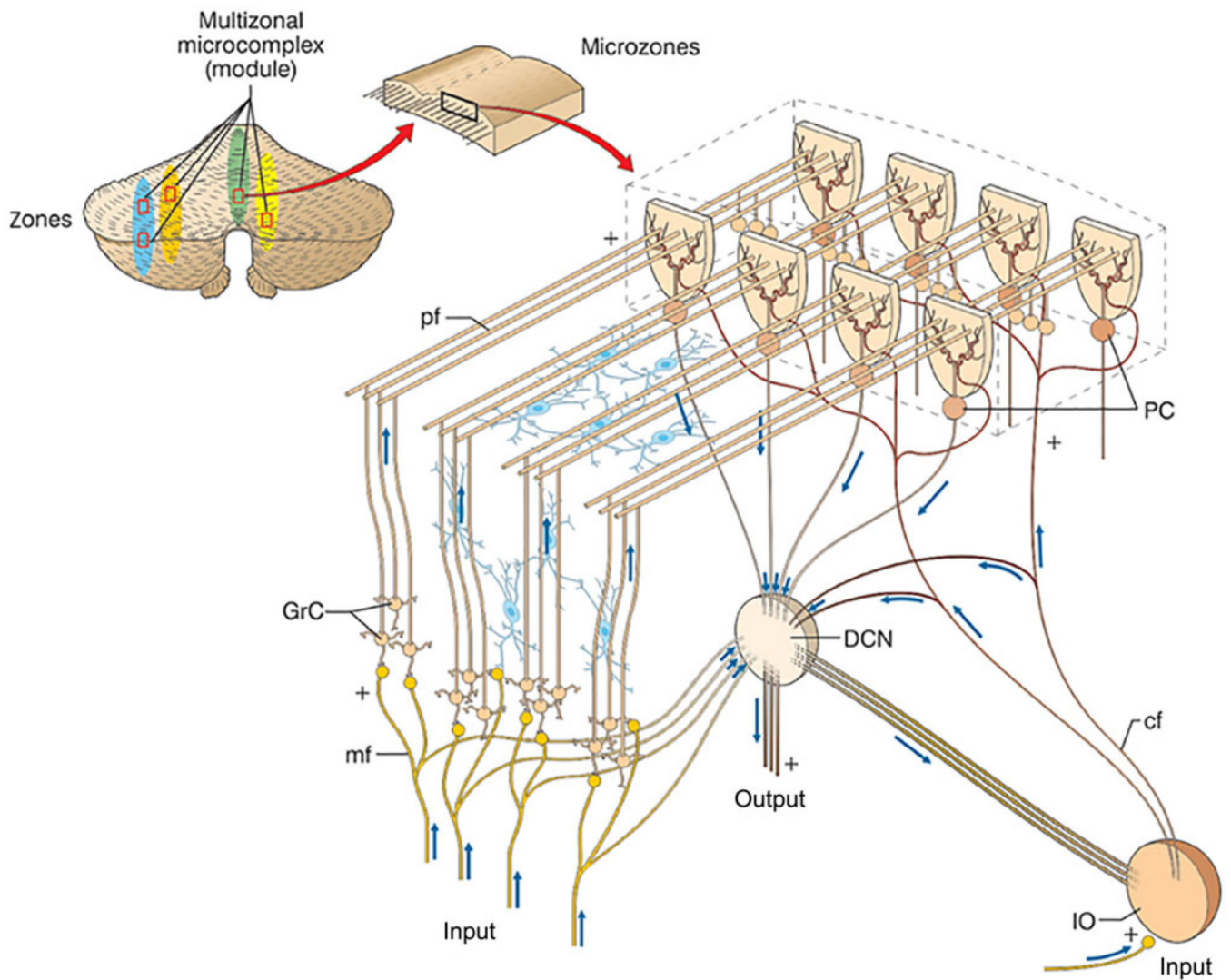
Figure 2.

Microcircuitry of the cerebellum. mf, mossy fibers (arising from the pontine nuclei); pf, parallel fiber; GrC, granule cell; PC, Purkinje cell; DCN, deep cerebellar nucleus; IO, inferior olive; cf, climbing fiber. Arrows indicate the direction of information flow, while + denotes excitatory synapses. Reproduced, with permission, from Ref. 169.
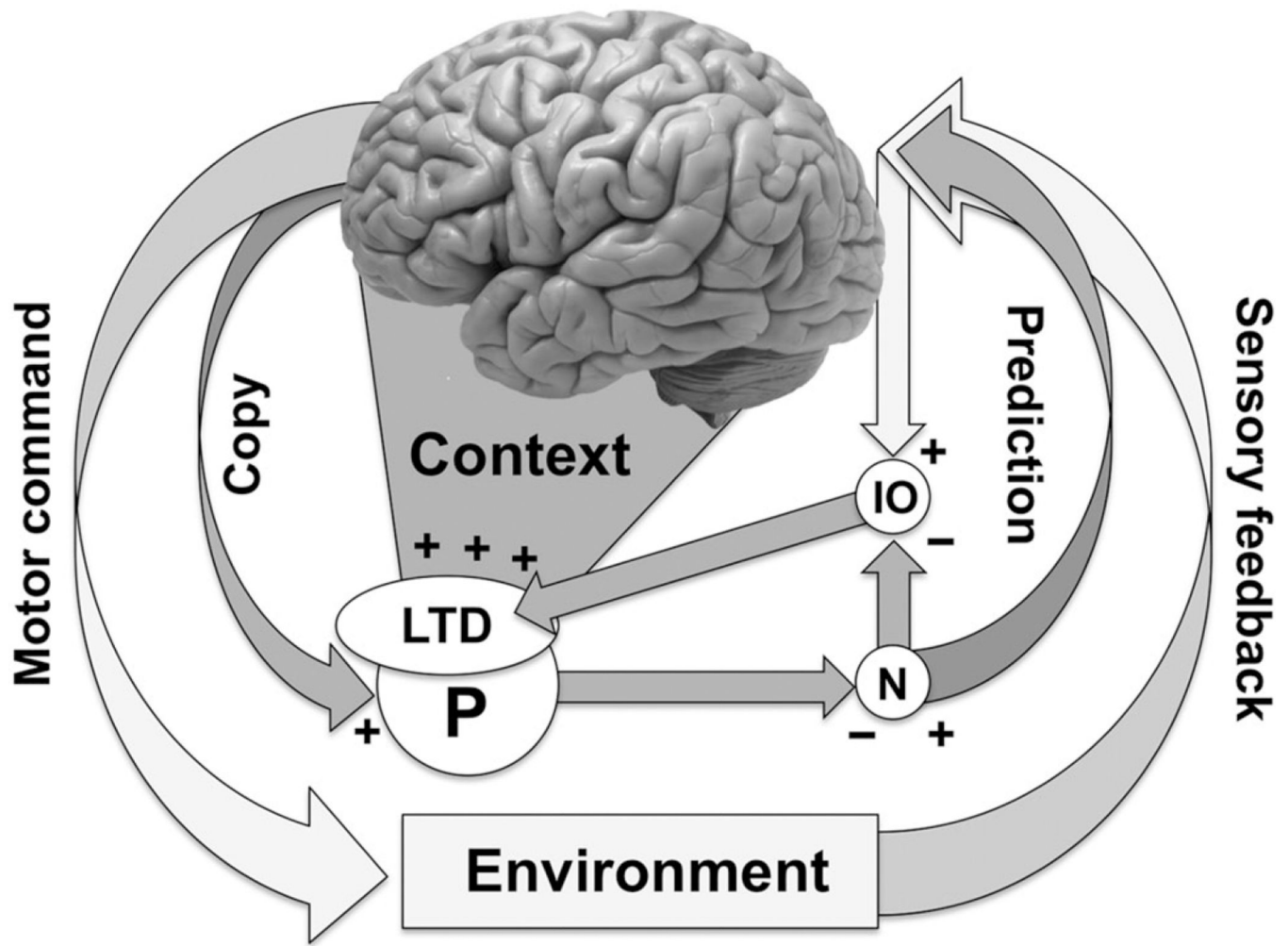
Figure 3.
The external control loop running through feedback from the environment (light gray arrows) and a cerebellar internal model (dark gray arrows). P, Purkinje cell; N, cerebellar nuclear cell; IO, inferior olive cell; LTD, long-term depression. + indicates excitatory, and − indicates inhibitory synapses. Context refers to multimodal information about the current state of the organism and its environment, relayed, like the copy of a motor command, via the mossy fiber pathway. The brain image was adapted with permission from www.flickr.com/photos/flamephoenix1991/8376271918.
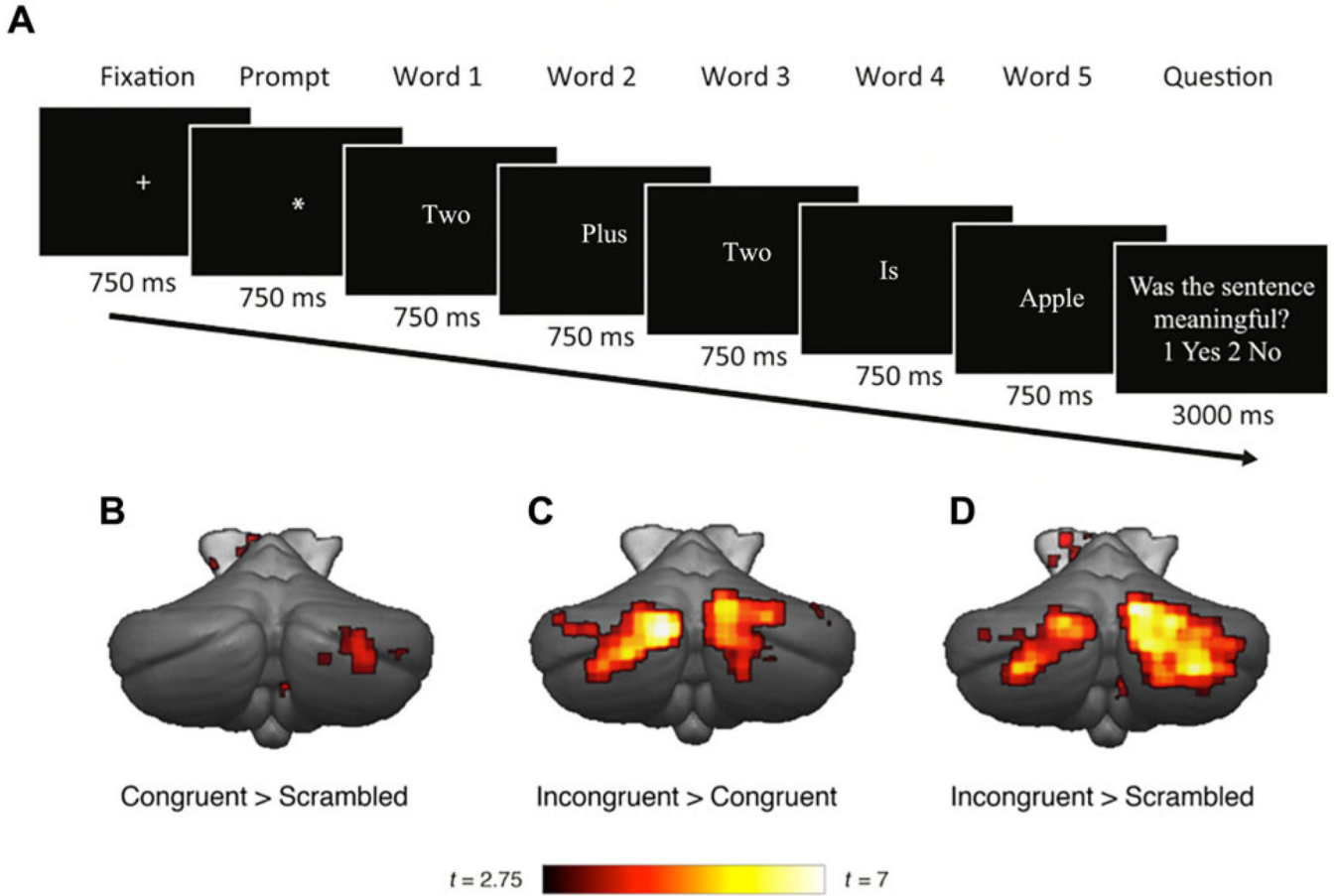
Figure 4.
(A) Task structure from the fMRI experiment in Ref. 145. In the *congruent* condition, sequentially presented words formed sentences where the target word was highly predictable (e.g., "two plus two is four"). In the *incongruent* condition, a strong contextual prediction was violated by an incongruent terminal word (e.g., "two plus two is apple."). In the *scrambled* condition, the initial four words did not establish a context for a grammatical sentence (e.g., "fast in clock plane"), and thus the target word was not predictable (e.g., "through"). (B) Cerebellar activations related to predictability. (C and D) Cerebellar activations related to prediction error. Adapted, with permission, from Ref. 145.