Elsevier Editorial System(tm) for Brain and Language
Manuscript Draft

Manuscript Number: BRLN-14-258R2

Title: The Cerebellum, Internal Models and Prediction in 'Non-motor' Aspects of Language: A Critical Review

Article Type: SI: Cerebellum

Keywords: language; prediction; cerebellum; internal models; associative learning; priming; default-mode network

Corresponding Author: Dr. Georgios P.D. Argyropoulos, PhD

Corresponding Author's Institution: Institute of Child Health, University College London

First Author: Georgios P.D. Argyropoulos, PhD

Order of Authors: Georgios P.D. Argyropoulos, PhD

Manuscript Region of Origin: UNITED KINGDOM

Abstract: The emergence of studies on cerebellar contributions in 'non-motor' aspects of predictive language processing has long been awaited by researchers investigating the neural foundations of language and cognition. Despite i) progress in research implicating the cerebellum in language processing, ii) the widely-accepted nature of the uniform, multi-modal computation that the cerebellum implements in the form of internal models, as well as iii) the long tradition of psycholinguistic studies addressing prediction mechanisms, research directly addressing cerebellar contributions to 'non-motor' predictive language processing has only surfaced in the last five years. This paper provides the first review of this novel field, along with a critical assessment of the studies conducted so far. While encouraging, the evidence for cerebellar involvement in 'non-motor' aspects of predictive language processing remains inconclusive under further scrutiny. Future directions are finally discussed with respect to outstanding questions in this novel field of research.

ORIGINAL MANUSCRIPT

The Cerebellum, Internal Models and Prediction in ‘Non-motor’ Aspects of Language: A Critical Review

Georgios P. D. Argyropoulos

Department of Psychology, Royal Holloway, University of London,

Egham, Surrey, TW20 0EX, UK

Correspondence:

Dr. Georgios P. D. Argyropoulos (gargyropoulos@gmail.com)

Abstract: The emergence of studies on cerebellar contributions in ‘non-motor’ aspects of predictive language processing has long been awaited by researchers investigating cerebellar involvement in higher cognition. Despite i) progress in research implicating the cerebellum in language processing, ii) the widely-accepted nature of the uniform, multi-modal computation that the cerebellum implements in the form of internal models, as well as iii) the long tradition of psycholinguistic studies addressing prediction mechanisms, research directly addressing cerebellar contributions to ‘non-motor’ predictive language processing has only surfaced in the last five years. This paper provides the first review of this novel field, along with a critical assessment of the studies conducted so far. While encouraging, the evidence for cerebellar involvement in ‘non-motor’ aspects of predictive language processing remains inconclusive under further scrutiny. Future directions are finally discussed with respect to outstanding questions in this novel field of research.

Keywords: language, prediction, cerebellum, internal models, associative learning, priming, default-mode network

1. *Introduction*

The cerebellum was traditionally seen as exclusively supporting the coordination of skilled, voluntary movement, gait, posture, balance, control of muscle tone, motor learning and articulation, with such involvement being reported for two centuries now. Readers are referred to Dow and Moruzzi (1958) for a thorough review of 19th century studies in the field, and to Manto et al. (2012) for a recent account. Over the last few decades, however, especially after the seminal work by the Leiners (see Leiner (2010) for a brief review), cumulative evidence has supported cerebellar involvement in a wide range of higher cognitive functions, e.g. memory, executive functions, visuospatial processing, emotional regulation, thought modulation, and, crucially, language (Mariën et al., 2014). In parallel, it has become increasingly clear that the cerebellum communicates in segregated anatomical loops with motor and prefrontal cortex (Strick, Dum & Fiez, 2009). Its cytoarchitectural uniformity supports the idea that its computations for motor control should guide hypotheses about its contributions in higher cognitive processes (Ramnani, 2006), including language (Ito, 2000a, 2008). However, research of relevance has, until very recently, been predominantly conducted in three poorly interfacing fields: i) work on cerebellar internal models in motor and non-motor aspects of behavior, without addressing their contribution in language; ii) research on cerebellar involvement in non-motor aspects of language processing, with no computational grounding; iii) studies conceptualizing predictive operations in language processing in terms of outputs of internal models, without addressing cerebellar circuitry as the most likely candidate for their implementation. This review will focus on studies of cerebellar involvement in the generation of semantic and phonological predictions above the lexical level.

2. *‘Cerebellum and ‘Prediction’, but no ‘Language’*

The contributions of the cerebellum in the automation of motor control have been well studied, and are consistent with the view that learning mechanisms store ‘motor memory’ through the diminishing synaptic efficacy between cerebellar neurons and their inputs (Marr, 1969; Albus, 1971). Control theoretic accounts suggest that such mechanisms allow cerebellar circuitry to acquire internal models that ultimately implement in an automatic fashion the movement-related processes initially established in the motor cortex (Ito, 1970, 1984; Wolpert, Miall, & Kawato, 1998; Ramnani, 2006). For instance, an internal (‘forward’) model of the arm’s dynamics, receives, as input, information on the current position and velocity of the arm, along with an ‘efference copy‘ of motor commands issued by the central nervous system, and outputs a prediction of the future position and velocity of the arm. Because of conduction delays in efferent and afferent pathways, the central nervous system is not immediately updated on changes in the peripheral motor system, and any recent commands issued may be yet to affect the musculature. Internal models are employed more rapidly, providing information about future properties of the controlled object, a fortiori in cases where accurate sensory feedback may be totally absent. This internal model ‘feedback’ allows the perceiver to rapidly interpret the perceptual signal and react accordingly, complete percepts received incompletely and/or under noise, and disambiguate in situations of uncertainty (Jordan & Wolpert, 2000). While internal models are conceivably located in all brain regions with synaptic plasticity that receive and send relevant information for their input and output (Kawato, 1999), a broad range of electrophysiological (Gilbert & Thach, 1977; Medina & Lisberger, 2009), imaging (Imamizu, Miyauchi, Tamada, Sasaki, Takino, Pütz, Yoshioka, & Kawato, 2000; Puttemans, Wenderoth, & Swinnen, 2005; Blakemore, Frith, & Wolpert, 2001; Desmurget, Gréa, Grethe, Prablanc, Alexander, & Grafton, 2001), and clinical studies (Müller & Dichgans, 1994; Nowak, Timmann, & Hermsdorfer, 2007) have established the cerebellum as ‘the most likely site for forward models to be stored’ (Kawato, Kuroda, Imamizu, Nakano, Miyauchi, & Yoshioka, 2003, p. 171).

However, the cerebellum exhibits a combination of two striking properties that support the involvement of its internal models beyond motor control: namely, its ‘essentially uniform, monotonously repetitive architecture’ (Schmahmann, 2000, p. 206) and its massive connectivity with virtually all major subdivisions of the brain. In particular, axonal fiber-tracing studies have identified projections from a broad range of neocortical areas to the ponto-cerebellar system, and even further to specific cerebellar lobules. These lobules return projections to the very same cerebral cortical areas via the cerebellar nuclei and thalamus, hence operating by means of segregated anatomical loops. Characteristically, the primary motor cortex selectively communicates with cerebellar lobules HV, HVI, HVIIb, and HVIIIa (Middleton &

Strick, 2000; Kelly & Strick, 2003). Importantly, though, a substantial range of prefrontal areas send inputs to the pontine nuclei, which, along with the inferior olive, form the two major sources of input to the cerebellum. These prefrontal areas span from area 10 through to posterior regions of area 8, crucially including area 45B in the rostral bank of the inferior limb of the arcuate sulcus (Schmahmann & Pandya, 1997). Prefrontal neurons provide inputs to Purkinje cells in lobule VIIa and Crura I and II in lobule HVIIa (Kelly & Strick, 2003). This has encouraged the idea that '[i]f closed-loop circuits reflect a general rule, then all of the areas of cerebral cortex that project to the cerebellum are the target of cerebellar output' (Strick, Dum, & Fiez, 2009, p. 422). Furthermore, lobule HVIIa boasts unparalleled computational power: it comprises nearly 50% of the cerebellar volume (Balsters et al., 2010), while the human cerebellum itself contains more neurons than all the rest of the nervous system (see Leiner (2010) for references). Moreover, recent work employing intrinsic functional connectivity in humans has demonstrated that HVIIa Crus I/II can be further subdivided on the basis of functional connectivity with anterior, ventromedial, and dorsolateral regions in the prefrontal cortex (Krienen & Buckner, 2009; Buckner, Krienen, Castellanos, Diaz, & Yeo, 2011). Since the uniformity of cellular organization across the cerebellar cortex implies identity in the computations performed (e.g. Bloedel, 1992; Schmahmann, 1997), the same forms of plasticity might support the automation of cognitive processes, and knowledge on cerebellar motor learning has indeed started to guide the study of cerebellar contributions to cognition (Ramnani, 2006; figure 1 below).

Associative learning represents a fundamental mechanism by which cerebellar internal models are seen to work. Classical conditioning, for instance, provides the most basic form of associative memory formation, and the cerebellum has long been established as a fundamental site (e.g. Thompson, Bao, Chen, Cipriano, Grethe, Kim, Thompson, Tracy, Weninger, & Krupa, 1997; Christian & Thompson, 2005) in both comparative (e.g. McCormick & Thompson, 1984) and human clinical studies (e.g. Daum, Ackermann, Schugens, Reimold, Dichgans, & Birbaumer, 1993; Timmann, Kolb, Baier, Rijntjes, Mueller, Diener, & Weiller, 1996). Crucially, the cerebellum is involved in non-motor aspects of associative learning. In a series of studies, cerebellar patients acquired associations between colors and numerals by trial-and-error. In comparison with normal controls, patients were significantly slower in learning the correct associations, and were impaired in recognizing them later. Control conditions established that this learning deficit could not be reduced to patients' motor impairment (Drepper, Timmann, Kolb, & Diener, 1999; Timmann, Drepper, Maschke, Kolb, Böring, Thilmann, & Diener, 2002; Timmann, Drepper, Calabrese, Bürgerhoff, Maschke, Kolb, Daum, & Diener, 2004). Moreover, recent fMRI work has shown that HVIIa Crura I/II are involved in the acquisition and employment of first- and second-order rules. In these studies, subjects acquired arbitrary associations by trial-and-error of visual stimuli with manual responses (first-order rules) or with instructions on the selection of a first-order rule (second-order rules). With the employment of delayed response tasks, activity time-locked to the onset of the rule-related cue was disambiguated from that for motor responses or visual feedback (Balsters & Ramnani, 2008, 2011; Balsters, Whelan, Robertson, & Ramnani, 2013). These findings corroborate early proposals that the cerebellum 'predicts' and 'prepares' the internal conditions required for sensory, motor, autonomic, memory-related, attention-related, affective, or linguistic operations, by acquiring the 'predictive relationships among temporally ordered multidimensional sequences of exogenously derived […] and endogenously derived […] neural activities' (Courchesne & Allen, 1997, p. 2). More recently, contributions of cerebellar internal models to multi-modal associative learning have been thoroughly examined in two reviews (Timmann, Drepper, Frings, Maschke, Richter, Gerwig, & Kolb, 2010; Bellebaum & Daum, 2011). Quite importantly, though, no considerations were present in either of those on the ways in which cerebellar internal models could contribute to non-motor aspects of linguistic associative learning and processing.

[Please insert figure 1 here]

3. *'Cerebellum and Language', but no 'Prediction'*

In parallel, studies on the cerebellar contributions in language have started to flourish over the last few decades. Although still often marginalized in neurobiological models of language processing (e.g. Hickok & Poeppel, 2007; Friederici, 2012), the 'ongoing enigma' of the linguistic cerebellum has been attracting significant attention (Mariën et al., 2014). Indeed, there has been cumulating clinical evidence for even a 'lateralized linguistic cerebellum' (Mariën, Engelborghs, Fabbro, & De Deyn, 2001), and meta-analyses of imaging studies locate activations related to language processing in right hemispheric cerebellar regions (E, Chen, Ho, & Desmond, 2014;

Stoodley & Schmahmann, 2009). The discussion of the cerebellum and language in toto is beyond the scope of this paper. For an up-to-date account of the evidence, the reader is encouraged to consult De Smet, Paquier, Verhoeven, and Mariën (2013), as well as papers in this special issue. Even so, the discussion of the way in which cerebellar internal models may be involved in language processing has remained poor, if not utterly absent, until very recently.

4. *'Language and Prediction', but no 'Cerebellum'*

Turning to prediction mechanisms in language processing, these are anything but novel a concept in studies of psycholinguistics and the neurobiology of language. For decades now, research has shown that predictions at the sentence level modulate speech perception and production (Lieberman, 1963) and accelerate syntactic processing (e.g. Staub & Clifton, 2006; see also discussion in Sturt & Lombardo, 2005).

Unsurprisingly, then, prediction in language processing is often conceptualized to operate in the form of internal model outputs. These are used in work on auditory feedback for speech production (Perkell, Matthies, Lane, Guenther, Wilhelms-Tricarico, Wozniak, & Guiod, 1997; Golfinopoulos, Tourville, Bohland, Ghosh, Nieto-Castanon, & Guenther, 2011; Golfinopoulos, Tourville, & Guenther, 2010), providing rapid information on the predicted auditory consequences of articulatory gestures well before the later-arriving sensory feedback (Heinks-Maldonado, Nagarajan, & Houde, 2006; Tourville, Reilly, & Guenther, 2008), or even in the absence of overt articulation (Tian & Poeppel, 2010). The N400, an event-related brain potential response, has also been discussed within the context of violated predictions generated by internal models (Hosemann, Herrmann, Steinbach, Bornkessel-Schlesewsky, & Schlesewsky, 2013).

In their seminal paper, Pickering and Garrod (2007) advanced the idea that internal models are employed in language processing in a multi-level fashion, extending beyond acoustic perception. Language perception and comprehension were proposed to dynamically combine the outputs generated by an 'input analysis system' with predictions of internal models providing an estimate of the next state that the input analysis system will enter (figure 2 below). Whenever the prediction is strong and the input noisy, the internal model exerts strong influence to the 'input analysis system'. In that way, the internal model ensures noise-resistant, rapid perception and comprehension of utterances. It should be noted that Pickering and Garrod's (2007) work draws from literature discussing the covert, imitative involvement of action production mechanisms in the efficient perception of actions performed by conspecifics (Wilson & Knoblich, 2005). Whether internal models in language comprehension recruit language generation mechanisms is beyond the scope of this paper, and remains an outstanding question (see section 7 below).

[Please insert figure 2 here]

5. Cerebellum, Prediction and Language

Given the well-established nature of cerebellar contributions in motor control (Manto et al., 2012), studies of speech production and perception have demonstrably synthesized insight from these three strands of research. Recent proposals emphasize the importance of the cerebellum and the basal ganglia in detecting temporal regularities and generating predictions during speech processing- hence alleviating the workload of fronto-temporal speech processing networks (Kotz & Schwartze, 2010). Accounts of language production also suggest that cerebellar internal models operate at the articulatory level (Hickok, 2012). Superior paravermal regions, for instance, may encode feedforward programs for the production of syllables (Ghosh, Tourville, & Guenther, 2009). Stuttering has been discussed within the context of deficient cerebellar internal models in generating a motor prediction error (Golfinopoulos, Tourville, & Guenther, 2010). Similarly, right cerebellar lobule HVI may undertake the temporal processing of segmental properties and left HVI may process prosodic, melodic properties (Callan et al., 2007). More recent work implicates the cerebellum in perceptual enhancement of acoustic properties of the linguistic signal. An fMRI study employed a word recognition task including acoustic stimuli from severely distorted speech. Improvements in the perception of such stimuli modulated cerebellar activity in four distinct cerebellar regions, one crucially being the right HVIIa Crus I. Activation in this lobule functionally correlated with cerebral regions that encompassed portions of the left angular and left temporal gyri. The findings supported the idea that regions within the left temporal and parietal cortex and the right Crus I (potentially along with lobules V/VI), participate in a functional network for achieving adaptive plasticity in speech perception. Discrepancies between the actual distorted acoustic speech input and the predicted acoustic input for a lexical item were discussed as engaging cerebellar-dependent supervised learning mechanisms (Guediche, Holt, Laurent, Lim, & Fiez, 2014).

The evidence accumulated so far has already encouraged discussions on the extension of the contributions of cerebellar internal models to non-motor aspects of language processing (Ito, 2000a, 2008; Argyropoulos, 2008, 2009). For instance, Argyropoulos (2009) argued that the cerebellum can store associative memory traces of contiguous linguistic events in sentence comprehension, such as the instance of processing the phonological, syntactic, and semantic properties of word ‘x’ and subsequently those of word ‘y’. In brief, cerebellar internal models would receive a copy of the input processed at each stage by a fronto-temporal ‘input analysis system’ (Pickering & Garrod, 2007) along with the operations performed on it, via the cerebro-ponto-cerebellar pathway. The internal models would then output a prediction of the next stage that the system would enter, via the ventrolateral neodentate, the parvocellular red nucleus and the thalamus back to the prefrontal cortex. Any discrepancies between the two would be conveyed as error signals back to the neocerebellar cortex through climbing fibers from the inferior olive. By long-term depression of parallel fiber-to-Purkinje cell synapses, these errors would train the internal model for accurate predictions. In cases where input analysis is conducted in noisy conditions and/or strong neocerebellar predictions are transmitted, neocerebellar output would override and/or bypass the output of cortico-cortical processing. With its massive computational power (see above), lobule HVIIa may successfully undertake such a Herculean task. Figure 3 provides a schematic illustration of phonological predictions generated in cerebro-cerebellar circuitry-conceivably, though, cerebellar circuitry would undertake the acquisition of temporal regularities and the implementation of predictions at different levels (e.g. phonology, syntax, semantics). The description of interactions among processes at different levels is beyond the scope of this paper. While the specialization of distinct lobular regions in different levels of prediction remains an outstanding question in the field (see section 7 below), a series of studies have followed these proposals, providing evidence for the possibility of cerebellar internal models to transmit predictions on upcoming ‘non-motor’ information in sentence processing.

[Please insert figure 3 here]

5.1. Cerebellar involvement in phrasal predictions

In the first cerebellar TMS study to use a language task, Argyropoulos (2011a) addressed cerebellar contributions to such ‘non-motor’ aspects of predictive language processing. At a minimal level, the capacity of the neocerebellum to bias and/or pre-empt the prefrontal cortex for upcoming linguistic input was predicted to manifest itself in phrasal associative, and not semantic categorical lexical priming. On the one hand, phrasal associative priming pertains to the probability that one word may call to mind a second one, and is based on the temporal contiguity and predictability of items in discourse. Pairs of lexical morphemes in idiomatic phrases provide a case in point, such as ‘gift-horse’, ‘skeletons-closet’. On the other hand, semantic categorical relatedness reflects the taxonomic relations between paradigmatic co-exemplars and the overlap in featural descriptions of two words, e.g. ‘tree’-‘bush’, ‘bee’-‘grasshopper’ (Thompson-Schill, Kurtz, & Gabrieli, 1998; Hutchison, 2003).

A lexical decision task assessed noun-to-noun priming sizes (quantified as differences in milliseconds between decision latencies for related and unrelated lexical pairs) before and after rTMS of the right neocerebellar vermis and of a control site in healthy native English speakers, as yielded by two different types of pairs: phrasal associates (e.g. ‘gift’-‘horse’), and pairs of subordinate and superordinate terms of the same category (e.g. ‘penny’-‘coin’). TMS of the neocerebellar vermis selectively enhanced phrasal associative priming in the form of increased differences between response latencies for lexical targets preceded by an unrelated prime (e.g. ‘battery’-‘horse’) and those preceded by a phrasally associated one (e.g. ‘gift-‘horse’). No such effects occurred for categorically related items after TMS, and stimulation of the control site did not affect associative priming sizes. This finding was consistent with evidence that cerebellar damage impairs verbal fluency by affecting phonemic rule-based word production, yet sparing semantic rule-based performance (Leggio, Silveri, Petrosini, & Molinari, 2000). Encouragingly, research in Alzheimer's disease and semantic dementia has established that temporal lobe lesions induce disruptions in categorical semantic, but not in associative priming (Rogers & Friedman, 2008). Moreover, cerebellar patients have been shown to ably perform category learning tasks, unlike patients with Parkinson's disease (Maddox, Aparicio, Marchant, & Ivry, 2005). This finding also echoes speculations on the significance of the cerebellum in storing and generating prefabricated, rote-memorized, idiomatic sequences, as opposed to propositionally composed ones: '[w]e learn and can recite ''Jabberwocky'' as movement and not at all as language […] We rote- memorize something that has so little linguistic or logical connection among the elements that it is learned as a movement. We can listen to what we say in order to get at what we otherwise can't remember […] But it is not something we know. It is buried in a rote-learning movement sequence' (Thach, 1997, pp. 91-2).

5.2. Cerebellar involvement in semantic predictions

Arguably, though, word-to-word predictions need not occur exclusively at the phonological level. There is substantial evidence in the lexical priming literature for semantic associations reflecting the predictive linkage of concepts based on world knowledge, as in 'instrument-action' pairs ('broom'-'sweep'), 'script relations' ('theatre'-'play'), 'locative relations' ('beach'-'house'), 'compositional relations' ('brick'-'house'; Hutchison, 2003; Jones & Estes, 2012). According to the theory on event schemata, stereotypical knowledge about the world is organized into units, including information about sequences of objects, situations, events, or states. Schemata are not the products of inferential manipulation of declaratively encoded representations. They operate online and below the level of awareness in sentence comprehension upon encountering the relevant concepts (e.g. the association of a dog with the agent-role in a biting event). In both dual-stage and constraint-based models of sentence comprehension, mechanisms of prediction (McRae, Hare, Elman, & Ferretti, 2005) and schema transmission (Ferreira, 2003; Townsend & Bever, 2001) employ these semantic associations in order to assign probable thematic (θ)-roles (Chomsky, 1981) and form a meaning-form hypothesis.

Suggestively, there is ample evidence from imaging, neurostimulation, and clinical studies in support of the involvement of the cerebellum in word generation tasks, crucially verb-to-noun generation. In a seminal PET study, subjects read aloud, repeated, passively read, or listened to nouns, or generated semantically appropriate verbs in response (e.g. ‘eat’ for ‘cake’). Blood flow changes occurred in paravermal regions when subjects read aloud or repeated nouns, as compared to when they viewed or listened to them. However, activation in the right lateral cerebellum was found for overt verb generation, but not for noun repetition/reading. This was difficult to account for on a motor basis. It was equally enigmatic though with respect to the involvement of the cerebellum in computations underlying verb generation (Petersen, Fox, Posner, Mintun, & Raichle, 1989). Recent fMRI studies provide similar results. Comparing verb generation and verb reading, Frings, Dimitrova, Schorn, Elles, Hein-Kropp, Gizewski, Diener and Timmann (2006) found activations in the right cerebellar lobule HVI and HVIIa Crus I as a measure of verb generation; again, these were lateral from the paravermal activation of lobule VI, which was associated with speech articulation. Fiez, Petersen, Cheney, and Raichle (1992) examined an English-speaking patient with a large right cerebellar infarct, who generated inappropriate responses in a number of word-generation tasks, involving verb generation. The patient’s responses were inappropriate, but remained categorically related (e.g. ‘small’, instead of take or ‘swallow’, in response to ‘pill’). This could not be attributed to overall cognitive impairment, as the patient’s performance on tests of memory, intelligence, ‘frontal function’, and language skills was excellent. This supports the idea that cerebellar damage leaves semantic networks intact (Fiez & Raichle, 1997). In another study, patients performed poorly in generating appropriate verbs for nouns, but selected the correct verb for a noun from a list of alternative responses, suggesting that semantic/syntactic representations were preserved. They were also able to produce appropriate subordinate term-responses to superordinate terms, suggesting that ‘[t]he right posterolateral cerebellum may be more involved in associative semantics than in categorical semantics’ (Gebhart, Petersen, & Thach, 2002). This pattern of impairments is qualitatively different from the direct disruptions in semantic networks associated with temporal cortical pathology in semantic dementia and Alzheimer’s disease (e.g. Rogers & Friedman, 2008). Studies on ‘associative’ and ‘semantic’ priming in Broca’s and Wernicke’s aphasia also show no impairments in automatically accessing lexical and/or semantic information (Blumstein, Milberg, & Shrier, 1982; Milberg & Blumstein, 1981), suggesting that associative priming does not rely on these perisylvian structures. As already discussed, cerebellar patients ably perform category learning tasks, unlike patients with Parkinson's disease (Maddox et al., 2005). Moreover, in a study employing a card-sequencing task, cerebellar patients with left lesions showed selective impairments in script sequences based on pictorial material, while those with right lesions were only impaired in script sequences requiring verbal elaboration (Leggio, Tedesco, Chiricozzi, Clausi, Orsini, & Molinari, 2008). Likewise, cerebellar cathodal tDCS has facilitated the rate and consistency of responses in a verb generation task, as compared to sham and anodal stimulation. These facilitatory effects occur in the case of cerebral cortical tDCS following anodal stimulation. They were hence explained in terms of disinhibition of the left prefrontal cerebral cortex resulting from the inhibitory effect of cathodal tDCS on the overall inhibitory tone that the cerebellum exerts on the cerebral cortex (Pope & Miall, 2012). Finally, in a TMS study, right cerebellar rTMS reduced category switching in the form of reduced phonemic and semantic fluency (Arasanz, Staines, Roy, & Schweizer, 2012).

The involvement of neocerebellar circuitry in the generation of semantic predictions was first directly assessed in a cerebellar rTMS study (Argyropoulos, 2011b; Argyropoulos & Muggleton, 2013). A pair-wise lexical decision task used noun-primes, the semantic properties of which could be categorically related by synonymy (e.g. ‘theft’-‘stealing’) or thematically associated (denoting agents, patients, instruments or locations of actions) with their verb-targets (e.g. ‘chef’-‘cooking’). Four groups of subjects were employed: a group that underwent stimulation of a medial cerebellar site; a group that underwent stimulation of a right lateral cerebellar site corresponding to a region in right HVIIa Crus I involved in cognitive aspects of verb generation (Frings et al., 2006); two groups that completed the same session with no intervention of TMS, one with a 5’ break and another one without. Stimulation of this lateral cerebellar site selectively boosted associative priming. No effects appeared after medial cerebellar stimulation or no stimulation. Argyropoulos (2011b) and Argyropoulos and Muggleton (2013) discussed this effect as a disruption of inhibitory processes in generating predictions for associated events expressed by the verb-target, e.g. (cooking)´, upon processing the semantic properties of the noun-prime, e.g. (chef)´.

This study was soon followed up by Lesage, Morgan, Olson, Meyer, and Miall (2012; see also Lesage (2013) for stimuli used). The authors employed the ‘Visual

World' paradigm (Altmann & Kamide, 1999), assessing the onset latencies of listeners' saccadic eye movements towards images of objects that aurally presented sentences referenced. The authors used two types of sentences: 'predictive' sentences, where the object could be predicted among four alternatives (one target object and three distractor objects) on the basis of the semantic content of the verb (e.g. 'The man will sail the … boat/mountain/bird/car'), and control, 'non-predictive' sentences (e.g. 'The man will watch the … boat/mountain/bird/car'), where the verb was not selective for the target object. In pre-TMS conditions, priming occurred in the form of faster anticipatory responses in the predictive as compared to the non-predictive condition. Cerebellar rTMS significantly delayed such anticipatory responses. Stimulation did not change saccadic latencies in the non-predictive sentences, ruling out a general effect on language processing. It did not change eye movement kinematics either, thus ruling out disruptions in oculomotor control. The prediction deficit was moreover absent in two control groups (vertex stimulation and no stimulation), ruling out non-specific effects of stimulation.

More recently, two fMRI studies have attempted to address this issue. Both used sentence comprehension tasks manipulating cloze probabilities and hence the strength of 'semantic priming' (Moberget, Gullesen, Andersson, Ivry, & Endestad, 2014; Lesage, Hansen, & Miall, 2014). The task of the first study involved the presentation of a sequence of five centrally presented words. The authors manipulated cloze probability by varying the context provided by the sequence of the initial four words. In a 'Congruent' condition, the target word was highly predictable on the grounds of the preceding four (e.g., 'two plus two is four'). In an 'Incongruent' condition, a final word was also highly predictable, but that prediction was violated by the target word presented (e.g., '[the water] had frozen to cars'). In a 'Scrambled' condition, the first four words made a non-grammatical sentence stem (e.g., 'fast in clock plane'), rendering the target word unpredictable (e.g., 'through'). In the end of each presentation, participants assessed whether the sentence was meaningful or not by a button press. In their contrasts of interest, 'Congruent > Scrambled' yielded a cluster of activation across right HVIIa Crus I/II. 'Incongruent > Congruent' yielded an activation cluster across left HVIIa Crus I/II, IV, and medial VIIa Crus I and another across right HVIIa Crus I/II, medial VIIa Crus I/II and VI. 'Incongruent > Scrambled' showed activation in right HVIIa Crus I/II, VIIa Crus I/II, and VI, and left

HVIIa Crus I/II and VIIa Crus I. The authors reasoned that both the generation of a prediction for an upcoming word based on its sentential context as well as the violation of that prediction are consistent with the involvement of neocerebellar internal models in non-motor aspects of language processing (Moberget et al., 2014).

Using an event-related fMRI design, the authors of the second study demonstrated that sentence predictability modulated the BOLD signal amplitude in a right HVIIa. Three events were modelled per trial: a context sentence (e.g. 'Greg went home for Christmas dinner'); the stem of a second sentence (e.g. 'His mum always cooked a'); the end of the second sentence (e.g. 'turkey'). Subjects read the sentences and pressed a button indicating the plausibility of the outcome. Crucially, the stem event did not require a motor response. Predictability was used as a parametric modulator for context and stem events. The authors showed that 'a cluster in the right posterolateral cerebellum […] was modulated by the predictability of the stem independent of outcome' (Lesage, Hansen, & Miall, 2014; see Lesage (2012, pp. 194-5) for stimuli).

5.3. Cerebellar contributions to the routinization of language processing

As discussed above, cerebellar internal models have been held not only to generate predictions about the next state of a simulated process, but also to gradually undertake its automatic implementation. Interestingly, casual adult dialogical interaction exhibits phenomena of routinized language processing, with lexical items and syntactic constructions of re-occurring phonological and semantic properties (Pickering & Garrod, 2004). The significance of the cerebellum in automatizing motor repertoires has been supported by clinical (e.g. Lang & Bastian, 2002) and imaging evidence (e.g. Jueptner, Frith, Brooks, Frackowiak, & Passingham, 1997; Toni, Krams, Turner, & Passingham, 1998), extending recently to higher cognitive aspects of rule learning (e.g. Balsters & Ramnani, 2011). In a tDCS study, cerebellar stimulation impaired the practice-dependent improvement in performance in the Sternberg task irrespective of polarity. On the contrary, tDCS over the dorsolateral prefrontal cortex left it unchanged, markedly affecting verbal working memory per se. Cerebellar stimulation also left visual evoked potentials unchanged, thus excluding visual cortex involvement (Ferrucci, Marceglia, Vergari, Cogiamanian, Mrakic-Sposta, Mameli, Zago, Barbieri, & Priori, 2008).

Crucially, cerebellar impairments induce similar automatization deficits in language-related processes. In word generation tasks, cerebellar patients reduce their response latencies poorly across repeated blocks of trials in comparison with normal controls (Fiez et al., 1992; Gebhart, Petersen, & Thach, 2002). Further evidence has been recently provided by Argyropoulos, Kimiskidis, and Papagiannopoulos (2011). The authors recruited two groups of participants of native speakers of Modern Greek for a TMS study conducted in two sessions, with the same stimuli presented in the second session. One group received stimulation centered at a site overlying the neocerebellar vermis in the first session, and on a deeper, control site in the second. Another group received stimulation in these two sites in reverse order. The study employed a pairwise lexical decision task. Lexical decision latencies for the group that underwent stimulation of the neocerebellar vermis in the second session did not become any shorter after stimulation, in contrast to all other conditions. Such disruption could not be explained on the grounds of sensorimotor processes, since the second group, who received neocerebellar vermal stimulation in their first session, showed a significant reduction of their reaction times in that session after stimulation. However, the disruption only appeared after stimulation of the target site selectively for the first group, who were encountering letter-string pairs for the second time. The automatization deficit thus pertained in particular to processing pairs that had been already encountered.

6. Critical assessment

The evidence presented above for the involvement of the cerebellum in 'non-motor' aspects of predictive language processing is certainly consistent with the unitary, multi-modal nature of the contributions of cerebellar internal models. It appears also to smoothly follow from the sizeable volume of suggestive evidence that has preceded it. However, as this section will demonstrate, there are a large number of significant issues that minimize the strength of these findings and limit their interpretation.

6.1. TMS studies

Despite the significant advantages of cerebellar neurostimulation with TMS (Grimaldi et al., 2013) and tDCS (Grimaldi et al., 2014) over cerebellar patient and fMRI studies, there are methodological issues that are often hard to overcome. For instance, in TMS, optimal localization of the target region is meaningfully achieved by retrieving coordinates of the area with the highest level of activation during performance of the same task in a previously completed fMRI session by the very same subject. Alternatively, high-resolution structural images may be used to position the coil according to the subject's underlying anatomy. When these are unavailable, scalp-based measurements from clear external landmarks are often used, such as the vertex or the inion (Devlin & Watkins, 2007). Indeed, all four TMS studies discussed here (Argyropoulos, 2011a; Argyropoulos, Kimiskidis, & Papagiannopoulos, 2011; Argyropoulos & Muggleton, 2013; Lesage et al., 2012), used the inion as an external landmark, thus making precise localization of the lobular cortex stimulated impossible. However, all four studies demonstrate further weaknesses that interact with the aforementioned limitation and compromise the inferences that can be drawn.

6.1.1. Stimulating the neocerebellar vermis

The putative target in the study of Argyropoulos (2011a) and Argyropoulos, Kimiskidis, and Papagiannopoulos (2011) was a right superior posterior vermal site. Their external landmark coordinates (1 cm below the inion, 1 cm laterally to right) targeted the medial (VIIa) Crus I/II. In order to estimate the depth of the site and thus the possibility of its successful stimulation, the authors recruited a volunteer whose brain image was already registered with a TMS-MRI co-registration system. The site corresponded to the right VIIa. Encouragingly, the superior posterior vermis is one of the lobules closest to the TMS coil (Miall & Christensen, 2004), and its stimulation has induced behavioral effects with high spatial precision (Hashimoto & Ohtsuka, 1995).

However, the anatomical connectivity of the vermis with the frontal lobes remains unclear (e.g. Kelly & Strick (2003) on medial VIIa-BA 46 connectivity; but also Coffman, Dum, & Strick (2011) on vermal-motor cortical connectivity). On the other hand, resting-state functional connectivity studies demonstrate that the vermis and the medial regions of the Crura contain a broad range of network nodes, the most posteromedial of which belongs to the dorsal attention network (Buckner et al., 2011). In another study, a cluster spanning across Crus II with the medial portions close to the vermis showed connectivity with the left executive control network

(Habas, Kamdar, Nguyen, Prater, Beckmann, Menon, & Greicius, 2009). More recently, Halko, Farzan, Eldaief, Schmahmann, and Pascual-Leone (2014) applied cerebellar rTMS guided by subject-specific connectivity to evaluate the relevance of connections between cerebral and cerebellar hubs belonging to different functional networks. One of their stimulation sites was lobule VII, and they demonstrated that its stimulation influences the cerebral dorsal attention system.

Equally enigmatic is the nature of vermal contributions in language. On the one hand, vermal lobule VII is involved in oculomotor control (e.g. Hashimoto & Ohtsuka, 1995) and also supports emotional processing (E, Chen, Ho, & Desmond, 2014; Stoodley & Schmahmann, 2009). On the other hand, superior posterior vermal activations are found in studies of language processing that are hard to reduce to motor effects (Desmond & Fiez, 1998; Desmond, Gabrieli, & Glover, 1998). Vermal tumor resection in children is associated with agrammatism (Riva & Giorgi, 2000). Voxel-based morphometric studies also report correlations between vermal grey matter and working memory measures (Ding, Qin, Jiang, Zhang, & Yu, 2012), and other studies find increased activity in the vermis after lexical training (Raboyeau, Marie, Balduyck, Gros, Démonet, & Cardebat, 2004). Vermal atrophy is also the most widely-cited cerebellar abnormality in schizophrenia (Picard, Amado, Mouchet-Mages, Olié, & Krebs, 2008): indeed, lexical priming studies on schizophrenic patients show greater priming for associated-only word pairs than for pairs related only semantically or both semantically and associatively (Nestor, Valdman, Niznikiewicz, Spencer, McCarley, & Shenton, 2006).

Evidently, though, the use of external landmarks for coil positioning could not exclude the probability that adjacent lobular cortices of VI, VIIb and VIII were also stimulated. While stimulation of right paravermal compartments is traditionally achieved by placing the coil 2 cm laterally to the right from the inion (e.g. Miall & Christensen, 2004) instead of 1cm, the scalp coordinates used here for targeting the right neocerebellar vermis may concomitantly stimulate portions of the right superior paravermis (e.g. Hashimoto & Ohtsuka, 1995). As discussed above, these lobules are reciprocally connected with the motor cortex in a segregated fashion (Kelly & Strick, 2003). While the second priming type used (semantically related and unrelated pairs) provided sufficient conditions for controlling for motor effects, the interpretation of the main effect here would be different. It can only be speculated that the selective effect on associative priming would pertain to cerebellar involvement in ‘silent/covert speech’ (e.g. Ackermann, Wildgruber, & Grodd, 1998). In an fMRI study, activation within the same region was found during silent recitation of the names of the months of the year (Ackermann, Mathiak, & Ivry, 2004), i.e. a routinized sequence of lexical morphemes. Such ‘inner speech’ has been held to provide a common platform for a broad range of cognitive functions implicating the cerebellum (Ackermann, Mathiak, & Riecker, 2007). This account would echo discussions whereby formulaic utterances are ‘buried in a rote-learning movement sequence’ engaging the cerebellum (Thach, 1997, pp. 91-2). In that way, it could only be speculated that the abnormalities in predictive functions were owed to impairments in the covert employment of the language production circuit (Pickering & Garrod, 2007).

For the study in Argyropoulos (2011a), in particular, a series of other weaknesses would include the low number of subjects and items per condition. Moreover, the within-subjects design involved participants being exposed to the same items twice (stimulation of control and target site counterbalanced across subjects). This may have introduced confounds with the effects that cerebellar stimulation has on practice-induced facilitation in performance (Ferrucci et al., 2008; Argyropoulos, Kimiskidis, & Papagiannopoulos, 2011). Also, the study in Argyropoulos, Kimiskidis and Papagiannopoulos (2011) failed to replicate the selective effects of neocerebellar vermal stimulation on associative priming. This could be attributed to the unavailability of published word association norms in Modern Greek for the construction of the stimulus set. On the other hand, the main finding in Argyropoulos, Kimiskidis, and Papagiannopoulos (2011) on the effects of neocerebellar vermal stimulation on the disruption of practice-induced accelerations of lexical decisions when participants encountered the same stimuli for the second time was not present in Argyropoulos (2011a). This could be due to the higher number of participants employed in that study.

6.1.2. Stimulating the neocerebellar hemispheres

Similarly, in Argyropoulos & Muggleton (2013), the target stimulation site was identified by external landmarks. The aim was to stimulate the peak coordinates of the activation cluster yielded in Frings et al. (2006) when comparing verb generation in inner speech with verb reading in inner speech (right HVIIa Crus I). Its coordinates were again registered and converted into scalp coordinates using a TMS-MRI co-registration system in one volunteer, and corresponded to 10 cm laterally to the right from the inion. The significant distance from that landmark would only increase the inter-subjective variability and the uncertainty on the cerebellar lobule stimulated. Moreover, the significant depth of the underlying tissue may have made it impossible to successfully induce an effect in at least a subset of participants with the figure-of-eight coil used (Hardwick, Lesage, & Miall, 2014), despite the high amplitude of stimulation. Another concern is that significant associative priming was only observed in the second phase of the experimental session across the four groups (control TMS, target TMS, no TMS 1, no TMS 2). While the associative boost after stimulation of this lateral site was significantly larger than that for the other groups and was also absent from the semantic categorical priming type, a concrete interpretation of the effect remains elusive. It would suggest that the processes affected were not automatic in nature and may reflect TMS effects on strategically employing semantic associations.

Similar issues in identifying the exact cerebellar lobular cortex underlying the stimulation site apply in Lesage et al*.* (2012), given that external landmarks were used again. However, the authors successfully stimulated the deeper hemispheric regions of the neocerebellum by employing a double-cone coil (Hardwick, Lesage, & Miall, 2014). As compared to figure-of-eight coils, though, focality of stimulation is more limited, and the possibility of a spread of the effect in neighbouring areas is increased. Suggestively, the authors used a set of coordinates (1 cm below the inion and 3 cm laterally to the right) that have been traditionally employed to induce cerebellar inhibition and in return facilitation in MEPs evoked by single pulse TMS of the contralateral primary motor cortex (e.g. Oliveri, Koch, Torriero, & Caltagirone, 2005). These coordinates have also been used by the same laboratory in recent assessments of successfully modulating motor cortical excitability with cerebellar stimulation by different coil types (Hardwick, Lesage, & Miall, 2014). They have also been employed in their previous TMS investigations to aim at the hand area of the ipsilateral cerebellar cortex, with its stimulation affecting cortico-cerebellar projections, changing motor cortical excitatability in the contralateral hand area (Miall, Christensen, Cain, & Stanley, 2007). Previous studies have demonstrated that this target site ‘affects the hand area of motor cortex, consistent with activation of lateral cerebellar cortex, probably in lobules V and VI' (Miall & King, 2007, p. 576). Therefore, despite the fact that in Lesage et al*.* (2012) stimulation with the same coordinates was 'directed towards Crus II' (Grimaldi et al., 2013, p. 133), motor-projecting cerebellar lobules HV and HVI were clearly also stimulated.

However, this limitation further interacts with a weakness in the construction of the stimulus set of the task. The authors acknowledge that they did not distinguish action-related from non-action related verbs, and were thus unable to test for any specific effects of action verbs in cerebellar motor-projecting lobules. They entertain the possibility that these are more active when processing action-related verbs. They reason, however, that, if such mechanisms were indeed present, they would not suffice to explain their results, given that the impairment was specific to sentences with predictive verbs, irrespective of any action-related semantics. The certainty with which such an interpretation can be dismissed comes into question when the stimuli (Lesage, 2013, p. 189) used in this study (Lesage et al., 2012) are examined under further scrutiny. The verbs in the 'predictive' condition score higher in concreteness, imageability, and meaningfulness metrics (e.g. Coltheart, 1981; Brysbaert, Warriner, & Kuperman, 2014) than those in the 'non-predictive' (control) condition. This is illustrated in table 1 below. The cerebellar lobules HV and HVI that were concomitantly stimulated form a segregated loop with the motor cortex (Kelly & Strick, 2003), and ample evidence associates action-related semantic processing with activity in motor structures used to implement those actions. For instance, processing speed for leg-related words is affected by TMS over the leg area of the left primary motor cortex, while performance remains unaffected when the left arm area or the right leg area is stimulated (Pulvermüller, Hauk, Nikulin, & Ilmoniemi, 2005). Similarly, fMRI studies show activations in passive reading of verbs that denote actions executed by different effectors are somatotopically organised in the primary motor cortex (e.g. Hauk, Johnsrude, & Pulvermüller, 2004). Another study has demonstrated that the MEP size in each effector muscle is only affected when listening to sentences containing actions related to that effector (Buccino, Riggio, Melli, Binkofski, Gallese, & Rizzolatti, 2005). Moreover, cathodal tDCS on the left motor cortex in healthy participants has reduced success rates in the acquisition of action-related words, in comparison with anodal or sham stimulation, tDCS over the prefrontal cortex, and learning of object-related words (Liuzzi et al., 2010).

Unsurprisingly, imagined movement also engages the cerebellum (e.g. Hanakawa et al., 2008). Some first suggestive evidence for the modulation of processing action-related verbs by cerebellar TMS can be found in Oliveri et al. (2009).

[Please inset table 1 about here]

What if stimulation was to selectively affect HVIIa Crus I/II, which communicates with the contralateral dorsolateral prefrontal cortex (Kelly & Strick, 2003)? The confounds in the design of the stimuli between sentence 'predictiveness' on the one hand and verb imageability, concreteness, and meaningfulness on the other would still not support the conclusion that cerebellar TMS disrupted cerebellar predictive processing. This is because of the modulation of activity in the dorsolateral prefrontal cortex by such lexical semantic properties. Concrete as compared to abstract concepts elicit greater activity in the left dorsolateral prefrontal cortex (Roxbury, McMahon, & Copland, 2014). Stimulation of the left dorsolateral prefrontal cortex also preferentially affects verbs compared to nouns (Shapiro, Pascual-Leone, Mottaghy, Gangitano, & Caramazza, 2001; Cappa, Sandrini, Rossini, Sosta, & Miniussi, 2002). This supports the interpretation that the functional link with motor cortical hand or leg regions may be mediated by the dorsolateral prefrontal cortex, much as the ventrolateral prefrontal cortex mediates the connection between speech perception and the motor cortical mouth region (Watkins & Paus, 2004): rTMS of the left dorsolateral prefrontal cortex has decreased naming latencies for verbs relative to its right homologue and sham stimulation, leaving latencies for object naming unaffected (Cappa et al., 2002). In another experiment, participants inflected nouns and verbs of a natural language as well as their pseudoword counterparts. Stimulation over the left dorsolateral prefrontal cortex affected reaction times selectively for verbs and not nouns (Shapiro et al., 2001). Therefore, the selective effect that cerebellar TMS had on 'predictive' sentences may be explained by the fact that the properties of the verbs in these sentences engaged to a significantly larger extent action-related semantic processing in the motor and/or dorsolateral prefrontal cortex than the more abstract and less imageable and meaningful verbs in 'non-predictive' sentences did. Stimulation of motor- or prefrontal-projecting cerebellar lobules may have thus disrupted this engagement.

6.2. fMRI studies

Turning to the two fMRI studies recently reported (Moberget et al., 2014; Lesage et al., 2014), a concern in the design of both is the confound between processing difficulty and ‘predictability’. As it shall be argued below, the activations in the particular regions of HVIIa Crus I/II that the authors report are probably yielded by differences in generic processing demands and not predictability per se.

6.2.1. Events: predictive or predicted?

To begin with, a question for both fMRI studies would pertain to the relevance of the event the onset of which hemodynamic activity was time-locked to. Indeed, the activations that both studies report pertain to activity time-locked to the onset of a stimulus (word or sentence) that could be predicted to a smaller or larger extent by the preceding context. In other words, the event of interest seems to be the occurrence of a predictable or unpredictable word or sentence, and not the occurrence of a context that invites the generation of predictions over subsequent input.

Moreover, in Moberget et al. (2014) the activity at the onset of the final word was not sufficiently disambiguated from and was susceptible to contamination by that for response preparation and implementation. Their sentence types did not only differ with respect to the predictability of the last word, but also with respect to the demands in response preparation. The activations they report in the contrast ‘Incongruent > Congruent’ provide a characteristic case in point. Indeed, the activations in this contrast fall within the region of lobule HVIIa Crus I/II that shows functional connectivity with Krienen and Buckner’s (2009) dorsolateral prefrontal cortical map and the frontoparietal control network (Buckner et al., 2011). However, this contrast need not reflect a violation of prediction in sentence comprehension. The difference in predictability is confounded with a number of discrepancies in other demands. Incongruent sentences here invite increased attention to and inhibitory control of an action being prepared (button key press to signal positive response to the question of meaningfulness) but needs to be revised (different button key press to signal negative response) upon encountering the last word; elevated demands for checking morpho-syntactic feature consistency and achieving semantic integration; elevated verbal working memory demands in reanalysing the semantic and syntactic structure of the sentence stem upon encountering the unpredictable target word.

In Lesage et al. (2014), such confounds are successfully avoided by sufficient jittering of the delay preceding the first context sentence as well as the second sentence stem. However, the HVIIa activations the authors report pertain to contrasts on activity time-locked to the onset of the stem of the second sentence. If this reflects the involvement of the cerebellum in generating predictions, it is quite puzzling why these activations are not present in comparisons on activity time-locked to the onset of the first sentence context, which invites stronger or weaker predictions on the content of the second sentence. The next section will make the case that these findings more plausibly reflect the modulation of hemodynamic activity by task difficulty, which is only confounded with predictability.

6.2.2. Networks: task-positive or task-negative?

Moberget et al*.* (2014) employed scrambled sentence stems as control contexts with minimal predictability for the final target word. In fact, the authors argued that, in studies similarly contrasting scrambled sentences with semantically and syntactically canonical sentences, 'the critical variable was the predictability of the presented words– with increased cerebellar involvement for predictable relative to unpredictable conditions.' (*ibid*, p. 2876). Yet scrambled sentences do not provide an unproblematic baseline if prediction per se is addressed in contrasting 'Congruent > Scrambled'. One major source of sentence comprehension difficulty is the interference of material that partially matches syntactic and semantic retrieval cues (see Glaser, Martin, Van Dyke, Hamilton, & Tan (2013) for discussion). Moreover, such anomalous sentences may engage the semantic integration process to a greater degree in participants' attempt to comprehend the sentence (Brown & Hagoort, 1993). For example, syntactic violations traditionally activate areas involved in syntactic processing, since structure building, agreement checking and other putative operations are disrupted, and hence extra attention is paid to these aspects (Kaan & Swaab, 2002). Greater activations often occur, for instance, in the left inferior frontal gyrus for violated sentences than for normal sentences (see discussion in Hagoort et al., 2009), and indeed scrambled sentences involve syntactic and semantic violations to a larger extent than sentences with no such violations. Cognitive control and conflict monitoring are arguably engaged more in scrambled sentences as compared to the other sentence types. Maintenance-related processes are also involved in keeping information accessible whenever the currently described events seize to be consistent with the global situation model under construction at each stage. With respect to sentence comprehension, then, scrambled sentences are arguably more taxing than congruent ones. This would mean that the ‘Congruent > Scrambled’ contrast would be reflecting not (just) ‘Predictable > Not predictable’, but also a more generic ‘Low Demands > High Demands’ discrepancy, owed to the larger demands in semantic and syntactic integration, cognitive control, conflict monitoring and more attempts for reanalysis made in the ‘Scrambled’ as compared to the ‘Congruent’ condition.

How then would a ‘Low Demands > High Demands’ contrast yield this activation in the posterolateral regions of HVIIa Crus I/II? To begin with, functional connectivity studies make clear that lobule (H)VIIa does not form a functionally unitary locus. While findings in different studies are not unequivocal, they show that there are distinct regions within this lobule that are functionally connected with different prefrontal cortical areas. In particular, they demonstrate that, apart from the well-established involvement of regions in (H)VIIa in the executive control network, there are distinct regions within this lobule that form crucial hubs of the default-mode network. These regions occupy a sizeable posterolateral portion of HVIIa Crus I/II (Krienen & Buckner, 2009; Buckner et al., 2011; Bernard et al., 2012; for discussion on the absence of such findings in Habas et al. (2009), see Buckner et al. (2011), p. 2324 and p. 2340). Hubs of the default-mode network consistently demonstrate increased activity during rest or low-demand tasks as opposed to high-demand tasks across a broad range of paradigms, and have been argued to support internal self-reflective thought. The suppression of this network in attention-demanding tasks is manifested in the form of ‘task-induced deactivations’, traditionally yielding significant activation in contrasts of ‘ Low Demands > High Demands’, given their larger negative BOLD signal amplitude in conditions posing higher demands (e.g. Mazoyer, Zago, Mellet, Bricogne, Etard, Houde, Crivello, Joliot, Petit, & Tzourio-Mazoyer, 2001; Binder, Frost, Hammeke, Bellgowan, Rao, & Cox, 1999; Shulman, Fiez, Corbetta, Buckner, Miezin, Raichle, & Petersen, 1997; McKiernan, Kaufman, Kucera-Thompson, & Binder, 2003; Gusnard & Raichle, 2001; Raichle, MacLeod, Snyder, Powers, Gusnard, & Shulman, 2001; Raichle & Gusnard, 2005).

There is indeed substantial evidence for the task-induced deactivation of these specific posterolateral regions of HVIIa Crus I/II, a fortiori in non-linguistic tasks and with no apparent demand for predictive processing. In one such study, participants were required to indicate whether a visually presented digit-symbol probe-pair was present or absent in an array of nine digit-symbol probe-pairs. Along with the deactivation of other major hubs of the default-mode network, the data also demonstrated strong deactivations in posterior regions of right lobule HVIIa Crus I/II (Rao, Motes, & Rypma, 2014). Elsewhere, Harrison et al. (2008) studied correlated activity fluctuations of the default mode network regions during three conditions: rest with eyes closed; a moral dilemma task; a Stroop task. They showed a striking uniformity in the anatomy of the default mode network across these conditions. Crucially, one of its hubs was a posterolateral region of HVIIa Crus I, showing activation in the resting state and moral dilemma, but deactivation in the Stroop task. Fransson (2005) also reports that one of the brain regions that correlated positively with other hubs of this network (precuneus / posterior cingulate cortex) during resting state and in a condition with eyes closed was the left and right posterolateral region of HVIIa Crus I. In another study (Fransson, 2006), one such node that showed significantly more activity during rest compared to a two-back working memory task was represented by a substantial cluster in posterolateral regions of the right HVIIa Crus I . Gusnard, Akbudak, Shulman, and Raichle (2001) examined hemodynamic activity related with judgments in an internally cued and an externally cued condition. Along with medial prefrontal cortical activations, the contrast of ‘Internally > Externally cued judgments’ yielded an activation in the posterolateral region of right HVIIa Crus I/II. Moreover, the HVIIa Crus I/II peak activation coordinates in these contrasts are markedly adjacent to those found in Krienen and Buckner (2009) as peak coordinates of functional connectivity with their medial prefrontal cortical map- a major hub of the default mode network. The involvement of these regions in the default-mode network need not negate their computational properties in implementing internal models. It may simply suggest that internal models in these regions support processes of the default-mode network. Table 2 below illustrates the adjacency of the peak activation coordinates from Moberget et al. (2014; ‘Congruent > Scrambled’) with those reported in studies on the engagement and disengagement of the default-mode network. Highly consistent with this interpretation is also the activation that Moberget et al. (2014) report in lobule IX for both ‘Congruent > Scrambled’ and ‘Incongruent > Scrambled’, as lobule IX is also part of the default mode network (Habas et al., 2009).

[Please insert table 2 here]

Likewise, in Lesage, Hansen, and Miall (2014), the parametric modulation by cloze probability in a right posterolateral region of HVIIa Crus II may not exclusively pertain to predictability. The confound with motor response preparation seen in Moberget et al. (2014) is certainly avoided. However, lower cloze probability still entails higher attentional and verbal working memory demands, as the memory trace of preceding linguistic information needs to be refreshed in order to check against a range of different completions. Indeed, the magnitude of deactivation of the default-mode network increases in a graded fashion in accordance with task load and error-proneness (McKiernan et al., 2003; Singh & Fawcett, 2008; Polli, Barton, Cain, Thakkar, Rauch, & Manoach, 2005; Persson, Lustig, Nelson, & Reuter-Lorenz, 2007; Kincses, Johansen-Berg, Tomassini, Bosnell, Matthews, & Beckmann, 2008; Esposito, Bertolino, Scarabino, Latorre, Blasi, Popolizio, Tedeschi, Cirillo, Goebel, & Di Salle, 2006). Hence the modulation of activity in this posterolateral region of right lobule HVIIa Crus II by linguistic predictability may simply reflect the modulation of the disengagement of the default mode network by task difficulty. This would be a by-product of predictive processing, which may instead be carried out by cerebral cortical regions.

6.3. *Directions for improvements*

The discussion above has made clear the need for a number of improvements in the methods and considerations on the different findings. For both TMS and fMRI studies, a substantial improvement would address the need for carefully designed stimulus sets, whereby distinct experimental conditions would be exclusively matched in all other aspects apart from those of interest, i.e. the predictability of upcoming stimuli. In any other case, modulation of cerebellar engagement by predictability may be confounded with other mechanisms which cerebellar circuitry may also support, such as those in processing action-related semantics (e.g. Lesage et al., 2012). Both TMS and fMRI studies should also take into account the functional connectivity of different regions within (H)VIIa Crus I/II with distinct prefrontal cortical areas. While a great part of this lobule is embedded in the executive control network, the existence of default-mode network hubs in certain posterolateral regions of HVIIa Crus I/II (e.g. Buckner et al., 2011; Bernard et al., 2012) suggests that these regions may show modulation by task difficulty as other hubs of this network in the cerebral cortex (e.g. Polli et al., 2007; Singh & Fawcett, 2008). This should be taken into account when interpreting the modulation of activity in lobule (H)VIIa by predictability in language processing, especially when activations occur within these posterolateral regions of lobule HVIIa (e.g. Moberget et al., 2014). Future studies would thus benefit from the use of further control conditions and/or tasks that concomitantly vary with respect to processing demands. These would help establish whether modulation of cerebellar activity by predictability is reduced to generic modulation by task load. Alternatively, future studies should show that linguistic predictability modulates activity in regions other than those that non-linguistic predictability does, or that different types of linguistic prediction (e.g. phrasal, semantic) yield activations in distinct cerebellar lobular regions. For TMS studies in particular, double-cone coils should be used to provide appreciable strength in the stimulation of the deeper cerebellar hemispheres (Lesage et al., 2012; Hardwick, Lesage, & Miall, 2014), as compared to flat figure-of-eight-shaped coils (Argyropoulos & Muggleton, 2013). On the other hand, the focality of the double-cone coil is limited as compared to that of a small figure-of-eight-shaped coil. Nevertheless, the stimulation of adjacent, motor-projecting lobules need not be an insurmountable obstacle, provided that the different experimental conditions are well-matched with respect to properties to which motor-projecting lobules may be sensitive (e.g. Lesage et al., 2012). While lobule VIIa also includes a vermal component, known as the ‘neocerebellar vermis’, the limited size of this portion and its adjacency to paravermal regions makes it difficult to stimulate in a selective fashion. This becomes an issue, when considering the involvement of paravermal regions in articulatory processes (Petersen et al., 1989; Frings et al., 2006). Irrespective of the coil shape and size, neuronavigated localization of the stimulation site (e.g. Halko et al., 2014) would contribute significantly towards identifying the particular lobular regions of (H)VIIa that receive the greatest stimulation.

7. *Outstanding questions*

However, aside from overcoming the aforementioned weaknesses in methods and the interpretations of the findings yielded so far, this newly emerging field of studies would benefit substantially from addressing a series of outstanding questions. Some of those are outlined below:

1. *What is the connectivity?* In control-theoretic accounts of cerebellar internal models, the anatomical evidence for the connectivity of specific cerebellar lobules with specific cerebral cortical areas is available for both motor control as well as higher cognition (Kelly & Strick, 2003). Discussions of cerebellar contributions to non-motor aspects of language processing have piggy-backed on the latter. Many authors have speculated that input to the cerebellum from Broca's area would provide an efference copy based on which the cerebellum would generate and transmit its predictions back to the prefrontal cortex (Ito, 2000a; Argyropoulos, 2009; Lesage et al., 2012). Functional connectivity studies (e.g. Buckner et al., 2011) are consistent with those ideas. However, apart from evidence on the cortico-pontine projection from area 45B (Schmahmann & Pandya, 1997) and the fact that such projections have massively expanded in humans (Ramnani et al., 2006), anatomical evidence for cerebello-dentate projections back to area 45B or for temporo-parietal cortical projections to the inferior olive remains poor (see Ito (2008) for some references).
2. *What about mirror neurons?* What is the relationship between cerebellar internal models and alleged cerebral cortical internal models that mirror neuron circuits implement (e.g. Iacoboni, 2005; Hurley, 2008)? Do cerebellar internal models output predictions by covertly employing the production system in perceiving conspecifics (e.g. Blakemore & Decety, 2001), like the alleged internal models in mirror-neurons? If so, what are the consequences for interpreting cerebellar involvement in language comprehension? Little attention has been paid to these questions so far (Miall, 2003; Strick, Dum, & Fiez, 2009).
3. *Is there a level-specific topography in the linguistic predictions of the cerebellum?* If the cerebellum is involved in predictive language processing after all, do different lobules or lobular regions generate predictions for different levels (phonology, semantics)? No study has so far identified the particular level at which these predictions occur. The effect of stimulating the posterior vermis/paravermis on phrasal predictions (Argyropoulos, 2011a) could in principle be contrasted to the effects of lateral cerebellar stimulation on semantic predictions (Lesage, 2012; Argyropoulos & Muggleton, 2013), but there are major inherent weaknesses in those studies to establish this.

4. *What about production?* The studies conducted so far have all addressed the involvement of cerebellar forward models in ‘non-motor’ aspects of prediction in language comprehension. This has been mainly driven by considerations of confounds with articulatory motor control that studying language generation would involve. However, prediction may involve both forward and inverse internal models at different levels of both language comprehension and generation (Pickering & Garrod, 2013). Indeed, the cerebellum may encode both forward and inverse internal models (Wolpert & Kawato, 1998). Can we then study cerebellar predictions in ‘non-motor’ aspects of language production?

8. Conclusion

Recent findings on the contributions of neocerebellar circuitry to the ‘non-motor’ aspects of predictive language processing are highly consistent with the involvement of the cerebellum in ‘higher cognitive’ aspects of behavior as well as with the multi-modal nature of this unitary cerebellar computation. However, a series of significant limitations in the few studies conducted so far raise questions about the interpretation of their findings. Addressing those weaknesses in the near future will encourage this newly emerging field to illustrate whether cerebellar internal models play a role in ‘non-motor’ aspects of predictive language processing. Above all, such research will enrich our understanding of the ‘modulatory’ role of the cerebellum in language processing from a computationally grounded perspective that integrates motor, cognitive and affective aspects of cerebellar function (Mariën et al., 2014).

References

Ackermann, H., Wildgruber, D., & Grodd, W. (1998). Does the cerebellum contribute to cognitive aspects of speech production? A functional MRI study in humans. *Neuroscience Letters*, *247*, 187-190.

Ackermann, H., Mathiak, K., & Ivry, R. B. (2004). Temporal organization of ―internal speech as a basis for cerebellar modulation of cognitive functions. *Behavioral and Cognitive Neuroscience Reviews*, *3*, 14-22.

Ackermann, H., Mathiak, K., & Riecker, A. (2007). The contribution of the cerebellum to speech production and speech perception: Clinical and functional imaging data. *Cerebellum*, *6*, 202–213.

Albus, J.S. (1971). A theory of cerebellar function. *Mathematical Biosciences*, *10*, 25– 61.

Altmann, G.T.M., & Kamide, Y. (1999). Incremental interpretation at verbs. *Cognition*, *73*, 247–264.

Arasanz, C.P., Staines, W.R., Roy, E.A., & Schweizer, T.A. (2012). The cerebellum and its role in word generation: a cTBS study. *Cortex*, *48*(6), 718-724.

Argyropoulos, G. P. (2008). The subcortical foundations of grammaticalization. In A. D. M. Smith, K. Smith, & R. Ferrer i Cancho (Eds.). *The evolution of language: Proceedings of the 7th international conference on the evolution of language* (pp. 10–17). Singapore: World Scientific.

Argyropoulos, G. P. (2009). Neocerebellar emulation in language processing. In K. Alter, M. Horne, M. Lindgren, M. Roll, & J. von Koss Torkildsen (Eds.). *Brain Talk: Discourse with and in the brain. Papers from the first Birgit Rausing language program conference in linguistics* (pp. 193–206). Lund: Lund University, Media Tryck.

Argyropoulos, G. P. (2011a). Cerebellar theta-burst stimulation selectively enhances lexical associative priming. *Cerebellum*, *10*(3), 540-50.

Argyropoulos, G.P. (2011b). *The neocerebellar Kalman filter linguistic processor: from grammaticalization to transcranial magnetic stimulation.* Unpublished doctoral dissertation. University of Edinburgh.

https://www.era.lib.ed.ac.uk/bitstream/1842/5694/2/Argyropoulos%202011.pdf

Argyropoulos, G. P., Kimiskidis, V., & Papagiannopoulos, S. (2011). Theta-burst stimulation of the right neocerebellar vermis selectively disrupts the practice-

induced acceleration of lexical decisions. *Behavioral Neuroscience*, *125*(5), 724-34.

Argyropoulos, G.P., & Muggleton, N. (2013). Effects of cerebellar stimulation on processing semantic associations. *Cerebellum*, *12*(1), 83-96.

Balsters, J.H., Cussans, E., Diedrichsen, J., Phillips, K.A., Preuss, T.M., Rilling, J.K., & Ramnani, N. (2010). Evolution of the cerebellar cortex: the selective expansion of prefrontal-projecting cerebellar lobules. *Neuroimage*, *49*, 2045-2052

Balsters, J. H., & Ramnani, N. (2011). Cerebellar Plasticity and Automation of Rule-Related Processing *Journal of Neuroscience*, *31*(6), 2305–2312.

Balsters, J. H., & Ramnani, N. (2008). Symbolic representations of action in the human cerebellum. *Neuroimage 43*(2), 388-398.

Balsters, J.H., Whelan, C.D., Robertson, I.H., & Ramnani, N. (2013). Cerebellum and cognition: evidence for the encoding of higher order rules. *Cerebral Cortex*, *23*(6), 433-43.

Bellebaum, C., & Daum, I. (2011). Mechanisms of cerebellar involvement in associative learning. *Cortex*, *47*(1), 128–36.

Bernard, J.A., Seidler, R.D., Hassevoort, K.M., Benson, B.L., Welsh, R.C., Wiggins, J.L., Jaeggi, S.M., Buschkuehl, M., Monk, C.S., Jonides, J., & Peltier, S.J. (2012). Resting state cortico-cerebellar functional connectivity networks: a comparison of anatomical and self-organizing map approaches. *Frontiers in Neuroanatomy*, *6*(31).

Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S., Rao, S. M., & Cox, R. W. (1999). Conceptual processing during the conscious resting state. A functional MRI study. *Journal of Cognitive Neuroscience*,*11*, 80–95.

Blakemore, S.-J., & Decety, J. (2001). From the perception of action to the understanding of intention. *Neuroscience*, *2*, 561-567.

Blakemore, S.-J., Frith, C. D., & Wolpert, D. M. (2001). The cerebellum is involved in predicting the sensory consequences of action. *NeuroReport*, *12*, 1879-84.

Bloedel, J.R. (1992). Functional heterogeneity with structural homogeneity - How does the cerebellum operate? *Behavioral and Brain Sciences*, *15*, 666–678.

Blumstein, S. E., Milberg, W. P., & Shrier, R. (1982). Semantic processing in aphasia: Evidence from an auditory lexical decision task. *Brain and Language*, *17*, 301–315.

Brown, C., & Hagoort, P. (1993). The processing nature of the N400: Evidence from

masked priming. *Journal of Cognitive Neuroscience*, *5*, 34 –44.

Brysbaert, M., Warriner, A.B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, *46*, 904-911.

Buccino, G., Riggio, L., Melli, G., Binkofski, F., Gallese, V., & Rizzolatti, G. (2005). Listening to action-related sentences modulates the activity of the motor system: a combined TMS and behavioral study. *Brain Research Cognitive Brain Research*, *24*, 355–63.

Buckner, R.L., Krienen, F.M., Castellanos, A., Diaz, J.C., & Yeo, B.T.T. (2011). The organization of the human cerebellum estimated by intrinsic functional connectivity *Journal of Neurophysiology, 106*, 2322–2345.

Callan, D.E., Kawato, M., Parsons, L., & Turner, R. (2007). Speech and song: the role of the cerebellum. *Cerebellum*, *6*, 321-327.

Canavan, A. G., Sprengelmeyer, R., Diener, H. C., & Hömberg, V. (1994). Conditional associative learning is impaired in cerebellar disease in humans. *Behavioral Neuroscience*, *108*, 475-85.

Cappa, S.F., Sandrini, M., Rossini, P.M., Sosta, K., & Miniussi, C. (2002). The role of the left frontal lobe in action naming: rTMS evidence. *Neurology*, *59*, 720–3.

Chomsky, N. (1981). *Lectures on government and binding.* Dordrecht: Foris.

Christian, K. M., & Thompson, R. F. (2005). Long-term storage of an associative memory trace in the cerebellum. *Behavioral Neuroscience*, *119*, 526-37.

Coffman, K., Dum, R. P., & Strick, P. L. (2011). Cerebellar vermis is a target of projections from the motor areas in the cerebral cortex. *Proceedings of the National Academy of Sciences*, *108*(38), 16068–73.

Coltheart, M. (1981). The MRC Psycholinguistic Database. *Quarterly Journal of Experimental Psychology*, *33*A, 497-505.

Courchesne, E., & Allen, G. (1997). Prediction and preparation, fundamental functions of the cerebellum. *Learning & Memory*, *4*(1), 1-35.

Daskalakis, Z.J., Paradiso, G.O., Christensen, B.K., Fitzgerald, P.B., Gunraj, C., & Chen, R. (2004). Exploring the connectivity between the cerebellum and motor cortex in humans. *Journal of Physiology*, *557*, 689–700.

Daum, I., Ackermann, H., Schugens, M.M., Reimold, C., Dichgans, J., & Birbaumer, N. (1993). The cerebellum and cognitive functions in humans. *Behavioral Neuroscience*, *107*(3), 411–19.

De Smet, H.J., Paquier, P., Verhoeven, J., & Mariën, P. (2013). The cerebellum: its role in language and related cognitive and affective functions. *Brain & Language*, *127*, 334-342.

Desmond, J. E., & Fiez, J. A. (1998). Neuroimaging studies of the cerebellum: Language, learning, and memory. *Trends in Cognitive Sciences*, *2*(9), 355-62.

Desmond, J. E., Gabrieli, J. D. E., & Glover, G. H. (1998). Dissociation of frontal and cerebellar activity in a cognitive task: evidence for a distinction between selection and research. *Neuroimage*, *7*, 368-76.

Desmurget, M., Gréa, H., Grethe, J. S., Prablanc, C., Alexander, G. E., & Grafton, S. T. (2001). Functional anatomy of nonvisual feedback loops during reaching: a positron emission tomography study. *Journal of Neuroscience*, *21*, 2919-2928.

Devlin, J. T., & Watkins, K. E. (2007). Stimulating language: insights from TMS. *Brain*, *130*(3), 610-622.

Ding, H., Qin, W., Jiang, T., Zhang, Y., & Yu, C. (2012). Volumetric variation in subregions of the cerebellum correlates with working memory performance. *Neuroscience Letters*, *508*(1), 47–51.

Dow, R. S., & Moruzzi, G. (1958). *The physiology and pathology of the cerebellum*. Minneapolis: University of Minnesota Press.

Drepper, J., Timmann, D., Kolb, F. P., & Diener, H. C. (1999). Non-motor associative learning in patients with isolated degenerative cerebellar disease. *Brain*, *122*, 87-97.

E, K.-H., Chen, S.-H. A., Ho, M.-H. R., & Desmond, J. E. (2014). A meta-analysis of cerebellar contributions to higher cognition from PET and fMRI studies. *Human Brain Mapping*, *35*(2), 593-615.

Esposito, F., Bertolino, A., Scarabino, T., Latorre, V., Blasi, G., Popolizio, T., Tedeschi, G., Cirillo, S., Goebel, R., & Di Salle, F. (2006). Independent component model of the default-mode brain function: Assessing the impact of active thinking. *Brain Research Bulletin, 70*, 263–269.

Fedorenko, E., Hsieh, P.-J., Nieto-Castañón, A., Whitfield-Gabrieli, S., & Kanwisher, N. (2010). New method for fMRI investigations of language: defining ROIs functionally in individual subjects. *Journal of Neurophysiology*, *104*(2), 1177–94.

Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology*, *47*, 164-203.

Ferrucci, R., Marceglia, S., Vergari, M., Cogiamanian, F., Mrakic-Sposta, S., Mameli,

F., Zago, S., Barbieri, S., & Priori, A. (2008). Cerebellar transcranial direct current stimulation impairs the practice-dependent proficiency increase in working memory. *Journal of Cognitive Neuroscience*, *20*(9), 1687-97.

Fiez, J. A., & Raichle, M. (1997). Linguistic processing. *International Review of Neurobiology, 41*, 233-54.

Fiez, J.A., Petersen, S.E., Cheney, M.K., & Raichle, M.E. (1992). Impaired nonmotor learning and error detection associated with cerebellar damage. A single case study. *Brain*, *115*, 155–78.

Fransson, P. (2005). Spontaneous low-frequency BOLD signal fluctuations: an fMRI investigation of the resting-state default mode of brain function hypothesis. *Human Brain Mapping, 26*, 15–29.

Fransson, P. (2006). How default is the default mode of brain function? further evidence from intrinsic BOLD signal fluctuations. *Neuropsychologia*, *44*, 2836–2845.

Friederici, A. D. (2012). The cortical language circuit: from auditory perception to sentence comprehension. *Trends in Cognitive Sciences*, *16*(5), 262–8.

Frings, M., Dimitrova, A., Schorn, C.F., Elles, H.-G., Hein-Kropp, C., Gizewski, E.R., Diener, H.C., & Timmann, D. (2006). Cerebellar involvement in verb generation: an fMRI study. *Neuroscience Letters*, *409*, 19–23.

Gebhart, A. L., Petersen, S. E., & Thach, W. T. (2002). Role of the posterolateral cerebellum in language. *Annals of the New York Academy of Science*, *978*, 318-33.

Ghosh S.S., Tourville, J.A., & Guenther F.H. (2009). A neuroimaging study of premotor lateralization and cerebellar involvement in the production of phonemes and syllables. *Journal of Speech, Language, and Hearing Research*, *51*(5), 1183-1202.

Gilbert, P.F. & Thach, W.T. (1977). Purkinje cell activity during motor learning. *Brain Research*, *128*, 309–328.

Glaser, Y.G., Martin, R.C., Van Dyke, J.A., Hamilton, A.C., & Tan, Y. (2013). Neural basis of semantic and syntactic interference in sentence comprehension. *Brain & Language*, *126*, 314–326.

Golfinopoulos, E., Tourville, J. a, Bohland, J. W., Ghosh, S. S., Nieto-Castanon, A., & Guenther, F. H. (2011). fMRI investigation of unexpected somatosensory feedback perturbation during speech. *Neuroimage*, *55*(3), 1324–38.

Golfinopoulos, E., Tourville, J., & Guenther, F. H. (2010). The integration of large-scale neural network modeling and functional brain imaging in speech motor control. *Neuroimage*, *52*(3), 862–74.

Greicius, M.D., & Menon, V. (2004). Default-mode activity during a passive sensor task: Uncoupled from deactivation but impacting activation. *Journal of Cognitive Neuroscience*, *16*, 1484–1492.

Grimaldi, G., Argyropoulos, G.P., Bastian, A., Cortes, M., Davis, N.J., Edwards, D., Ferrucci, R., Fregni, F., Galea, J.M., Hamada, M., Manto, M., Miall, R.C., Morales-Quezada, L., Pope, P.A., Priori, A., Rothwell, J., Tomlinson, S.P., & Celnik, P. (2014). Cerebellar transcranial direct current stimulation (ctDCS): a novel approach to understand cerebellar function in health and disease. *Neuroscientist* [epub ahead of print].

Grimaldi, G., Argyropoulos, G.P., Boehringer, A., Celnik, P., Edwards, M.J., Ferrucci, R., Galea, J.M., Groiss, S.J., Hiraoka, K., Kassavetis, P., Lesage, E., Manto, M., Miall, R.C., Priori, A., Sadnicka, A., Ugawa, Y., & Ziemann, U. (2013). Non-invasive cerebellar stimulation—a consensus paper. *Cerebellum*, *13*(1), 121-138.

Grush, R. (2004). The emulation theory of representation: motor control, imagery, and perception. *Behavioral and Brain Sciences*, *27*, 377-435.

Guediche, S., Holt, L. L., Laurent, P., Lim, S.-J., & Fiez, J. A. (2014). Evidence for cerebellar contributions to adaptive plasticity in speech perception. *Cerebral Cortex*. [Epub ahead of print].

Gusnard, D.A., & Raichle, M.E. (2001). Searching for a baseline: functional imaging and the resting human brain. *Nature Reviews Neuroscience, 2*, 685–694.

Gusnard, D.A., Akbudak, E., Shulman, G.L., & Raichle, M.E. (2001). Medial prefrontal cortex and self-referential mental activity: relation to a default mode of brain function. *Proceedings of the National Academic of Sciences*, 98(7), 4259–4264.

Habas, C., Kamdar, N., Nguyen, D., Prater, K., Beckmann, C. F., Menon, V., & Greicius, M. D. (2009). Distinct cerebellar contributions to intrinsic connectivity networks. *Journal of Neuroscience*, *29*(26), 8586–94.

Hagoort, P., Baggio, G., & Willems, R. M. (2009). Semantic unification. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences*, 4th ed. (pp. 819-836). Cambridge, MA: MIT Press.

Halko, M.A., Farzan, F., Eldaief, M.C., Schmahmann, J.D., & Pascual-Leone, A.

(2014). Intermittent theta-burst stimulation of the lateral cerebellum increases functional connectivity of the default network. *Journal of Neuroscience, 34*(36), 12049-12056.

Hanakawa, T., Dimyan, M.A., & Hallett, M. (2008). Motor planning, imagery, and execution in a distributed motor network: a time-course study with functional MRI. *Cerebral Cortex*, *18*, 2775–88.

Hardwick, R.M., Lesage, E., & Miall, R.C. (2014). Cerebellar transcranial magnetic stimulation: the role of coil geometry and tissue depth. *Brain Stimulation*, *7*, 643-649.

Harrison, B.J., Pujol, J., López-Solà, M., Hernández-Ribas, R., Deus, J., Ortiz, H., Soriano-Mas, C., Yücel, M., Pantelis, C., & Cardoner, N. (2008). Consistency and functional specialization in the default mode brain network. *Proceedings of the National Academic of Sciences*, *105*(28), 9781–9786.

Hashimoto, M., & Ohtsuka, K. (1995). Transcranial magnetic stimulation over the posterior cerebellum during visually guided saccades in man. *Brain*, *118*, 1185–1193.

Hauk, O., Johnsrude, I., & Pulvermuller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron*, *41*, 301–7.

Heinks-Maldonado, T. H., Nagarajan, S. S., & Houde, J. F. (2006). Magnetoencephalographic evidence for a precise forward model in speech production. *Neuroreport*, *17*(13), 1375–1379.

Hickok, G. (2012). Computational neuroanatomy of speech production. *Nature Reviews Neuroscience*, *13*, 135–145.

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, *8*, 393–402.

Hosemann, J., Herrmann, A., Steinbach, M., Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2013). Lexical prediction via forward models: N400 evidence from German Sign Language *Neuropsychologia*, *51*, 2224–2237.

Hurley, S. (2008). The shared circuits model (SCM): How control, mirroring, and simulation can enable imitation, deliberation, and mindreading. *Behavioral and Brain Sciences*, 31, 1- 58.

Hutchison, K. A. (2003). Is semantic priming due to association strength or feature overlap? A *micro*analytic review. *Psychonomic Bulletin & Review*, *10*(4), 785-813.

Iacoboni, M. (2008). The role of premotor cortex in speech perception: evidence from

fMRI and rTMS. *Journal of Physiology*, *102*(1-3), 31-34.

Imamizu, H, Miyauchi, S., Tamada, T., Sasaki, Y., Takino, R., Pütz, B., Yoshioka, T., & Kawato, M. (2000). Human cerebellar activity reflecting an acquired internal model of a new tool. *Nature*, *403*, 192-5.

Ito, M. (1970). Neurophysiological basis of the cerebellar motor control system. *International Journal of Neurology*, *7*, 162-76.

Ito, M. (1984). *The cerebellum and neural control.* New York: Raven Press.

Ito, M. (2000a). Neural control of cognition and language. In A. Marantz, Y. Miyashita, & W. O'Neil (Eds.), *Image, language, brain* (pp. 149-62). Cambridge, Massachusetts: MIT Press.

Ito, M. (2000b). Neurobiology: internal model visualized. *Nature*, *403*, 153-154.

Ito, M. (2008). Control of mental activities by internal models in the cerebellum. *Nature Reviews Neuroscience*, *9*, 304-13.

Jones, L.L., & Estes, Z. (2012). Lexical Priming: associative, semantic, and thematic influences on word recognition. In: J. Adelman (Ed.), *Visual Word Recognition, Vol II: Meaning and Context, Individuals and Development* (pp. 44-72). Psychology Press.

Jordan, M. I., & Wolpert, D. M. (2000). Computational motor control. In M. S. Gazzaniga, (Ed.), *The new cognitive neurosciences* (pp. 601-618). Cambridge, MA: MIT Press.

Jueptner, M., Frith, C.D., Brooks, D.J., Frackowiak, R.S.J., & Passingham, R.E. (1997). Anatomy of motor learning. II. Subcortical structures and learning by trial and error. *Journal of Neurophysiology*, *77*, 1325-37.

Kaan, E., & Swaab, T.Y. (2002). The brain circuitry of syntactic comprehension. *Trends in Neurosciences*, *6*(8), 350-356.

Kawato, M. (1999). Internal models for motor control and trajectory planning, *Current Opinion in Neurobiology*, *9*, 718-27.

Kawato, M., Kuroda, T., Imamizu, H., Nakano, E., Miyauchi, S., & Yoshioka, T. (2003). Internal forward models in the cerebellum: fMRI study on grip force and load force coupling. *Progress in Brain Research*, *142*, 171-88.

Kelly, R.M., & Strick, P.L. (2003). Cerebellar loops with motor cortex and prefrontal cortex of a nonhuman primate. *Journal of Neuroscience*, *23(23)*, 8432–44.

Kincses, Z.T., Johansen-Berg, H., Tomassini, V., Bosnell, R., Matthews, P.M., & Beckmann, C.F. (2008). Model-free characterization of brain functional networks

for motor sequence learning using fMRI. *Neuroimage*, *39*(4), 1950–1958.

Kotz, S.A., & Schwartze, M. (2010). Cortical speech processing unplugged: a timely subcortico-cortical framework. *Trends in Cognitive Sciences*, *14*, 392-399.

Krienen, F.M., & Buckner, R.L. (2009). Segregated fronto-cerebellar circuits revealed by intrinsic functional connectivity. *Cerebral cortex*, *19*, 2485-2497.

Lang, C. E., & Bastian, A. J. (2002). Cerebellar damage impairs automaticity of a recently practiced movement. *Journal of Neurophysiology*, *87*(3), 1336-47.

Leggio, M.G., Silveri, M.C., Petrosini, L., & Molinari, M. (2000). Phonological grouping is specifically affected in cerebellar patients: a verbal fluency study. *Journal of Neurology, Neurosurgery, and Psychiatry*, *69*, 102–6.

Leggio, M.G., Tedesco, A. M., Chiricozzi, F. R., Clausi, S., Orsini, A., & Molinari, M. (2008). Cognitive sequencing impairment in patients with focal or atrophic cerebellar damage. *Brain*, *131*, 1332-1343.

Leiner, H.C. (2010). Solving the mystery of the human cerebellum. *Neuropsychology Review*, *20*, 229-235.

Lesage, E. (2013). *On the cerebellum and language: neurostimulation and imaging studies.* Unpublished doctoral dissertation. University of Birmingham. http://etheses.bham.ac.uk/4926/1/Lesage14PhD.pdf

Lesage, E., Hansen, P.C., & Miall, C.C. (2014). Cerebellar BOLD response to linguistic stimuli is modulated by predictability. *Society for Neuroscience abstracts.*

Lesage, E., Morgan, B.E., Olson, A.C., Meyer, A.S., & Miall, R.C. (2012). Cerebellar rTMS disrupts predictive language processing, *Current Biology*, *22*(18), 794-795.

Lieberman, P. (1963). Some effects of semantic and grammatical context on the production and perception of speech. *Language and Speech*, *6*, 172-187.

Liuzzi, G., Freundlieb, N., Ridder, V., Hoppe, J., Heise, K, Zimerman, M., Dobel, C., Enriquez-Geppert, S., Gerloff, C., Zwitserlood, P., & Hummel, F.C. (2010). The involvement of the left motor cortex in learning of a novel action word lexicon. *Current Biology*, *20*(19), 1745-51.

Maddox, W.T., Aparicio, P., Marchant, N.L., & Ivry, R.B. (2005). Rule-based category learning is impaired in patients with Parkinson's disease but not in patients with cerebellar disorders. *Journal of Cognitive Neuroscience*, *17*(5), 707-23.

Manto, M., Bower, J.M., Conforto, A.B., Delgado-García, J.M., da Guarda, S.N., Gerwig, M., Habas, C., Hagura, N., Ivry, R.B., Mariën, P., Molinari, M., Naito, E., Nowak, D.A., Oulad Ben Taib, N., Pelisson, D., Tesche, C.D., Tilikete, C., &

Timmann, D. (2012). Consensus paper: roles of the cerebellum in motor control--the diversity of ideas on cerebellar involvement in movement. *Cerebellum*, *11*(2), 457-87.

Mariën, P., Ackermann, H., Adamaszek, M., Barwood, C.H.S., Beaton, A., Desmond, J., De Witte, E., Fawcett, A.J., Hertrich, I., Küper, M., Leggio, M., Marvel, C., Molinari, M., Murdoch, B.E., Nicolson, R.I., Schmahmann, J.D., Stoodley, C.J., Thürling, M., Timmann, D., Wouters, E., & Ziegler, W. (2014). Consensus paper: language and the cerebellum: an ongoing enigma. *Cerebellum*, *13*, 386–410.

Mariën, P., Engelborghs, S., Fabbro, F., & De Deyn, P. P. (2001). The lateralized linguistic cerebellum: a review and a new hypothesis. *Brain and Language*, *79*, 580-600.

Marr, D. (1969). A theory of cerebellar cortex. *Journal of Physiology*, *202*, 437– 470.

Mazoyer, B., Zago, L., Mellet, E., Bricogne, S., Etard, O., Houde, O., Crivello, F., Joliot, M., Petit, L., & Tzourio-Mazoyer, N. (2001). Cortical networks for working memory and executive functions sustain the conscious resting state in man. *Brain Research Bulletin*, *54*, 287–298.

McCormick, D. A, & Thompson, R. F. (1984). Cerebellum: essential involvement in the classically conditioned eyelid response. *Science*, *223*, 296-9.

McKiernan, K.A., Kaufman, J.N., Kucera-Thompson, J., & Binder, J.R. (2003). A parametric manipulation of factors affecting task-induced deactivation in functional neuroimaging. *Journal of Cognitive Neuroscience, 15*, 394–408.

McRae, K., Hare, M., Elman, J. L., & Ferretti, T. (2005). A basis for generating expectancies for verbs from nouns. *Memory and Cognition*, *33*(7), 1174-84.

Medina, J.F. & Lisberger, S.G. (2009). Encoding and decoding of learned smooth-pursuit eye movements in the floccular complex of the monkey cerebellum, *Journal of Neurophysiology*, *102*, 2039-2054.

Miall, R. C. (2003). Connecting mirror neurons and forward models. *NeuroReport*, *14*(16), 1-3.

Miall, R. C., & Christensen, L. O. D. (2004). The effect of rTMS over the cerebellum in normal human volunteers on peg-board movement performance. *Neuroscience Letters*, 371, 185-189.

Miall, R. C., Christensen, L. O. D., Cain, O., & Stanley, J. (2007). Disruption of state estimation in the human lateral cerebellum. *Public Library of Science Biology*, *5*(11), 2733-44.

Miall, R. C., & King, D. (2008). State estimation in the cerebellum. *Cerebellum*, *7*, 572-6.

Middleton, F.A. & Strick, P.L. (2000). Basal ganglia and cerebellar loops: motor and cognitive circuits. *Brain Research Reviews*, *31*, 236-250.

Middleton, F. A., & Strick, P. L. (2001). Cerebellar projections to the prefrontal cortex of the primate. *Journal of Neuroscience*, *21*(2), 700-12.

Milberg, W. P., & Blumstein, S. E. (1981). Lexical decision and aphasia: evidence for semantic processing. *Brain and Language*, *14*, 371–385.

Moberget, T., Gullesen, E.H., Andersson, S., Ivry, R.B., & Endestad, T. (2014). Generalized role for the cerebellum in encoding internal models: evidence from semantic processing. *Journal of Neuroscience*, *34*(8), 2871–2878.

Müller, F., & Dichgans, J. (1994). Dyscoordination of pinch and lift forces during grasp in patients with cerebellar lesions. *Experimental Brain Research*, *101*, 485-92.

Nestor, P., Valdman, O., Niznikiewicz, M., Spencer, K., McCarley, R., & Shenton, M. (2006). Word priming in schizophrenia: associational and semantic influences *Schizophrenia Research*, *82*(2), 139-42.

Nowak, D. A., Timmann, D., & Hermsdorfer, J. (2007). Dexterity in cerebellar agenesis. *Neuropsychology*, *45*, 696-703.

Oliveri, M., Koch, G., Torriero, S., & Caltagirone, C. (2005) Increased facilitation of the primary motor cortex following 1 Hz repetitive transcranial magnetic stimulation of the contralateral cerebellum in normal humans. *Neuroscience Letters*, *376*, 188–193.

Oliveri, M., Bonnì, S., Turriziani, P., Koch, G., Gerfo, E. L., Torriero, S., Vicario, C.M., Petrosini, L. & Caltagirone, C. (2009). Motor and linguistic linking of space and time in the cerebellum. *PloS One*, *4*(11), e7933.

Perkell, J. S., Matthies, M. L., Lane, H., Guenther, F. H., Wilhelms-Tricarico, R., Wozniak, J. & Guiod, P. (1997). Speech motor control: acoustic goals, saturation effects, auditory feedback and internal models. *Speech Communication*, *22*, 227-250.

Persson, J., Lustig, C., Nelson, J., & Reuter-Lorenz, P. (2007). Age differences in deactivation: a link to cognitive control? *Journal of Cognitive Neuroscience*, *19*, 1021–1032.

Petersen, S. E., Fox, P. T., Posner, M. L., Mintun, M., & Raichle, M. E. (1989).

Positron emission tomographic studies of the processing of single words. *Journal of Cognitive Neuroscience*, *1*, 153-70.

Picard, H., Amado, I., Mouchet-Mages, S., Olié, J.- P., & Krebs, M.-O. (2008). The role of the cerebellum in schizophrenia: an update of clinical, cognitive, and functional evidences. *Schizophrenia Bulletin*, *34*(1), 155-172.

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, *27*, 169-225.

Pickering, M. J., & Garrod, S. (2007). Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences*, *11*(3), 105-10.

Pickering, M. J., & Garrod, S. (2013). Forward models and their implications for production, comprehension, and dialogue. *Behavioral and Brain Sciences*, *36*(4), 377-392.

Poeppel, D., Emmorey, K., Hickok, G., & Pylkkänen, L. (2012). Towards a new neurobiology of language. *Journal of Neuroscience*, *32*,14125–14131.

Polli, F., Barton, J., Cain, M., Thakkar, K., Rauch, S., & Manoach, D. (2005). Rostral and dorsal anterior cingulate cortex make dissociable contributions during antisaccade error commission. *Proceedings of the National Academy of Science*, *102*, 15700–15705.

Pope, P., & Miall, R.C. (2012). Task-specific facilitation of cognition by cathodal transcranial direct current stimulation of the cerebellum. *Brain Stimulation*, *5*, 84-94.

Pulvermüller, F., Hauk, O., Nikulin, V.V., & Ilmoniemi, R.J. (2005). Functional links between motor and language systems. *European Journal of Neuroscience*, *21*, 793–7.

Puttemans, V., Wenderoth, N., & Swinnen, S.P. (2005). Changes in brain activation during the acquisition of a multifrequency bimanual coordination task: from the cognitive stage to advanced levels of automaticity. *Journal of Neuroscience, 25*, 4270–4278.

Raboyeau, G., Marie, N., Balduyck, S., Gros, H., Démonet, J. -F., & Cardebat, D. (2004). Lexical learning of the English language: a PET study in healthy French subjects. *Neuroimage*, *22*(4), 1808–18.

Raichle, M.E., & Gusnard, D.A. (2005). Intrinsic brain activity sets the stage for expression of motivated behaviour. *Journal of Comparative Neurology*, *493*(1), 167-176.

Raichle, M.E., MacLeod, A.M., Snyder, A.Z., Powers, W.J., Gusnard, D.A., & Shulman, G.L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences*, *98*, 676–682.

Ramnani, N. (2006). The primate cortico-cerebellar system: anatomy and function. *Nature Reviews Neuroscience*, *7*, 511–522.

Ramnani, N., Behrens, T.E., Johansen-Berg, H., Richter, M.C., Pinsk, M.A., Andersson, J.L., Rudebeck, P., Ciccarelli, O., Richter, W., Thompson, A.J., Gross, C.G., Robson, M.D., Kastner, S., & Matthews, P.M. The evolution of prefrontal inputs to the cortico-pontine system: diffusion imaging evidence from macaque monkeys and humans (2006). *Cerebral Cortex*, *16*(6), 811-8.

Rao, N.K., Motes, M.A., & Rypma, B. (2014). Investigating the neural bases for intra-subject cognitive efficiency changes using functional resonance imaging. *Frontiers in Human Neuroscience*, *8*, 840.

Riva, D., & Giorgi, C. (2000). The cerebellum contributes to higher functions: evidence from a series of children surgically treated for posterior fossa tumours. *Brain*, *123*, 1051-1061.

Rogers, S. L., & Friedman, R. B. (2008). The underlying mechanisms of semantic memory in Alzheimer’s disease and semantic dementia. *Neuropsychologia*, *46*, 12-21.

Roxbury, T., McMahon, K., & Copland, D.A. (2014). An fMRI study of concreteness effects in spoken word recognition. *Behavioral and Brain Functions*, *10*(34),1-14.

Schmahmann, J. D. (2000). The role of the cerebellum in affect and psychosis. *Journal of Neurolinguistics*, *13*, 189-214.

Schmahmann, J. D., & Pandya, D. N. (1997). The cerebrocerebellar system. *International Review of Neurobiology*, *41*, 31-60.

Shapiro, K.A., Pascual-Leone, A., Mottaghy, F.M., Gangitano, M., & Caramazza, A. (2001). Grammatical distinctions in the left frontal cortex. *Journal of Cognitive Neuroscience*, *13*, 713–20.

Shulman, G.L., Fiez, J.A., Corbetta, M., Buckner, R.L., Miezin, F.M., Raichle, M.E., & Petersen, S.E. (1997). Common blood flow changes across visual tasks: I. Increases in subcortical structures and cortex. *Journal of Cognitive Neuroscience*, *9*, 648–663.

Singh, K.D., & Fawcett, I.P. (2008). Transient and linearly graded deactivation of the human default-mode network by a visual detection task. *Neuroimage*, *41*,100-112.

Staub, A., & Clifton, C. Jr. (2006). Syntactic prediction in language comprehension: evidence from either…or. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(2), 425-436.

Stoodley, C.J., & Schmahmann, J.D. (2009). Functional topography in the human cerebellum: a meta-analysis of neuroimaging studies. *Neuroimage, 44*(2), 489-501.

Strick P.L., Dum, R.P., & Fiez, J.A. (2009). Cerebellum and nonmotor function. *Annual Review of Neuroscience*, *32*, 413–434.

Sturt, P., & Lombardo, V. (2005). Processing coordinated structures: incrementality and connectedness. *Cognitive Science*, *29*, 291-305.

Thach, W. T. (1997). Context-response linkage. *International Review of Neurobiology*, *41*, 599-611.

Thompson, R. F., Bao, S., Chen, L., Cipriano, B. D., Grethe, J. S., Kim, J. J., Thompson, J. K., Tracy, J. A., Weninger, M. S., & Krupa, D. J. (1997). Associative learning. *International Review of Neurobiology*, *41*, 152-89.

Thompson-Schill, S. L., Kurtz, K. J., & Gabrieli, J. D. E. (1998). Effects of semantic and associative relatedness on automatic priming. *Journal of Memory and Language*, *38*, 440-58.

Tian, X., & Poeppel, D. (2010). Mental imagery of speech and movement implicates the dynamics of internal forward models. *Frontiers in Psychology*, *1*, 1–23.

Timmann, D., Drepper, J., Calabrese, S., Bürgerhoff, K., Maschke, M., Kolb, F.P., Daum, I., & Diener, H. C. (2004). Use of sequence information in associative learning in control subjects and cerebellar patients. *Cerebellum*, 3, 75–82.

Timmann, D., Drepper, J., Frings, M., Maschke, M., Richter, S., Gerwig, M., & Kolb, F.P. (2010). The human cerebellum contributes to motor, emotional and cognitive associative learning. A review. *Cortex*. *46*, 845–57.

Timmann, D., Drepper, J., Maschke, M., Kolb, F.P., Böring, D., Thilmann, A.F., & Diener, H.C. (2002). Motor deficits cannot explain impaired cognitive associative learning in cerebellar patients. *Neuropsychologia*, *40*, 788–800.

Timmann, D., Kolb, F. P., Baier, C., Rijntjes, M., Mueller, S. P., Diener, H. C., & Weiller, C. (1996). Cerebellar activation during classical conditioning of the human flexion reflex: a PET study. *NeuroReport*, *7*, 2056-60.

Toni, I., Krams, M., Turner, R., & Passingham, R. E. (1998). The time course of changes during motor sequence learning: a whole-brain fMRI study. *Neuroimage*,

8, 50-61.

Tourville, J., Reilly, K. J., & Guenther, F. H. (2008). Neural mechanisms underlying auditory feedback control of speech. *Neuroimage*, *39*(3), 1429–43.

Townsend, D., & Bever, T. G. (2001). *Sentence comprehension: the integration of habits and rules.* Cambridge, Massachusetts: MIT Press.

Tucker, J., Harding, A. E., Jahanshahi, M., Nixon, P. D., Rushworth, M., Quinn, N. P., Thompson, P. D., & Passingham, R. E. (1996). Associative learning in patients with cerebellar ataxia. *Behavioral Neuroscience*, *110*, 1229-34.

Watkins, K., & Paus, T. (2004). Modulation of motor excitability during speech perception: the role of Broca's area. *Journal of Cognitive Neuroscience*, *16*(6), 978-987.

Wilson, M., & Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psychological Bulletin*, *131*(3), 460-473

Wolpert, D.M., & Kawato, M. (1998). Multiple paired forward and inverse models for motor control. *Neural Networks*, *11*, 1317–1329.

Wolpert, D.M., Miall, R.C., & Kawato, M. (1998). Internal models in the cerebellum. *Trends in Cognitive Sciences*, *2*(9), 338–47.

Figure Legends

Figure 1. Neural implementation of internal (forward) models in the cerebellum; left: an internal model implemented in cerebellar lobules HV,HVI,HVIIb,HVIII interacting with the motor cortex; right: an internal model implemented in cerebellar lobule (H)VIIa Crus I/II interacting with the prefrontal cortex. Its organization is explicitly based on the well-studied cerebro-cerebellar circuitry and corresponding control-theoretic accounts of motor control (left). A copy of a process involving the interaction of the prefrontal cortex with another brain area is sent via the pontine nuclei to its interconnected cerebellar lobules. Predictions generated from those are transmitted from the Purkinje cells of the cerebellar cortex via the ventrolateral neodentate nucleus and the thalamus back to those prefrontal areas. Predicted and actual consequences of the process copied by these cerebellar lobules are compared in the inferior olive. Any discrepancies between the two are fed via climbing fibers to the cerebellar cortex in the form of an error signal. Long-term depression is triggered at the parallel fiber-to-Purkinje cell synapses, updating the internal model. RN: red nucleus. Modified with permission from Macmillan Publishers Ltd: NATURE REVIEWS NEUROSCIENCE (Ramnani, N. (2006). The primate cortico-cerebellar system: anatomy and function. Nat. Rev. Neurosci., 7, 511–522), copyright (2006).

Figure 2. Schematic representation of an internal (forward) model involved in the comprehension of the sentence ‘Harry went out to fly his red flag’ for three different levels of prediction (Pickering & Garrod, 2007). This ‘emulator’ (Grush, 2004) is controlled by feedback from a Kalman filter, weighing predictions against analysis of

the speech input at each (arbitrarily defined here) step. In case of strong predictions and/or noisy inputs, there is ‘low Kalman gain’ (strong top-down influence on the analysis); in case of poor predictions and clear input, there is ‘high Kalman gain’ (strong bottom-up influence). For the sake of the illustration, the probabilistic nature of the predictions and the interaction among predictions at different levels are not discussed. The implementation of the internal model in the language production system will not be discussed here. Figure reproduced from Pickering & Garrod (2007, p. 108) with permission © 2007 Elsevier

Figure 3. Schematic illustration of the phonological aspects of lexical prediction generated by a cerebellar internal model. Its organization is explicitly based on the well-studied cerebro-cerebellar circuitry and corresponding control-theoretic accounts of motor control (figure 1 above). A copy of a process involving the interaction of the prefrontal cortex with another brain area (such as the temporo-parietal cortex) in the context of comprehending a sentence is sent via the pontine nuclei to its interconnected cerebellar lobules. Cerebellar predictions are transmitted from the Purkinje cells of the cerebellar cortex via the ventrolateral neodentate nucleus and the thalamus back to cortical areas. Predicted and actual consequences of the process are compared in the inferior olive. Any discrepancies between the two are fed via climbing fibers from the inferior olive to the cerebellar cortex in the form of an error signal. LTD (long-term depression) is triggered at the parallel fiber-to-Purkinje cell synapses, updating the internal model. For the sake of the illustration, the cortico-cortical interactions are assumed to proceed in a bottom-up fashion (see section 7 for more discussion). See also Ito(2008) for discussion of available evidence from anatomical connectivity studies supporting these models. Same

simplifications are made as in Pickering and Garrod (2007; figure 2 here).

Table Legends

Table 1. Concreteness, imageability, and meaningfulness ratings from Coltheart (1981) and Brysbaert, Warriner, and Kuperman (2014) for the available verbs used in the ‘non-predictive’ and ‘predictive’ sentences in Lesage et al. (2012). Verbs in ‘predictive’ sentences are more concrete, imageable, and meaningful as compared to those in ‘non-predictive’ sentences. Stimuli retrieved from Lesage (2013).

Table 2. Peak activation coordinates (MNI) in posterolateral regions of HVIIa Crus I/II for: the contrast ‘Congruent > Scrambled’ in Moberget et al. (2014); the functional connectivity study of Krienen and Buckner (2009); contrasts in a selection of fMRI studies on default-mode network. The coordinates of Rao, Motes, and Rypma (2014) are converted from Talairach to MNI space.

REVISED MANUSCRIPT

The Cerebellum, Internal Models and Prediction in ‘Non-motor’ Aspects of Language: A Critical Review

Georgios P. D. Argyropoulos

Developmental Neurosciences Programme

UCL Institute of Child Health

30 Guilford Street, London

WC1N 1EH

UK

Correspondence:

Dr. Georgios P. D. Argyropoulos (g.argyropoulos@ucl.ac.uk)

Abstract: The emergence of studies on cerebellar contributions in ‘non-motor’ aspects of predictive language processing has long been awaited by researchers investigating the neural foundations of language and cognition. Despite i) progress in research implicating the cerebellum in language processing, ii) the widely-accepted nature of the uniform, multi-modal computation that the cerebellum implements in the form of internal models, as well as iii) the long tradition of psycholinguistic studies addressing prediction mechanisms, research directly addressing cerebellar contributions to ‘non-motor’ predictive language processing has only surfaced in the last five years. This paper provides the first review of this novel field, along with a critical assessment of the studies conducted so far. While encouraging, the evidence for cerebellar involvement in ‘non-motor’ aspects of predictive language processing remains inconclusive under further scrutiny. Future directions are finally discussed with respect to outstanding questions in this novel field of research.

Keywords: language, prediction, cerebellum, internal models, associative learning, priming, default-mode network

1. *Introduction*

The cerebellum was traditionally seen as exclusively supporting the coordination of skilled, voluntary movement, gait, posture, balance, control of muscle tone, motor learning and articulation, with such involvement being reported for two centuries now. Readers are referred to Dow and Moruzzi (1958) for a thorough review of 19th century studies in the field, and to Manto et al. (2012) for a recent account. Over the last few decades, however, especially after the seminal work by the Leiners (see Leiner (2010) for a brief review), cumulative evidence has supported cerebellar involvement in a wide range of higher cognitive functions, e.g. memory, executive functions, visuospatial processing, emotional regulation, thought modulation, and, crucially, language (Mariën et al., 2014). In parallel, it has become increasingly clear that the cerebellum communicates in segregated anatomical loops with motor and prefrontal cortex (Strick, Dum & Fiez, 2009). Its cytoarchitectural uniformity supports the idea that its computations for motor control should guide hypotheses about its contributions in higher cognitive processes (Ramnani, 2006), including language (Ito, 2000a, 2008). However, research of relevance has, until very recently, been predominantly conducted in three poorly interfacing fields: i) work on cerebellar internal models in motor and non-motor aspects of behavior, without addressing their contribution in language; ii) research on cerebellar involvement in non-motor aspects of language processing, with no computational grounding; iii) studies conceptualizing predictive operations in language processing in terms of outputs of internal models, without addressing cerebellar circuitry as the most likely candidate for their implementation. This review will focus on studies of cerebellar involvement in the generation of semantic and phonological predictions above the lexical level in language comprehension. The interaction of predictive processes at different levels and the covert employment of language production mechanisms in comprehension (Pickering & Garrod, 2007) remain beyond the scope of this review.

2. *'Cerebellum and 'Prediction', but no 'Language'*

The contributions of the cerebellum in the automation of motor control have been well studied, and are consistent with the view that learning mechanisms store 'motor memory' through the diminishing synaptic efficacy between cerebellar neurons and their inputs (Marr, 1969; Albus, 1971). Control theoretic accounts suggest that such mechanisms allow cerebellar circuitry to acquire internal models that ultimately implement in an automatic fashion the movement-related processes initially established in the motor cortex (Ito, 1970, 1984; Wolpert, Miall, & Kawato, 1998; Ramnani, 2006). For instance, an internal ('forward') model of the arm's dynamics, receives, as input, information on the current position and velocity of the arm, along with an 'efference copy' of motor commands issued by the central nervous system, and outputs a prediction of the future position and velocity of the arm. Because of conduction delays in efferent and afferent pathways, the central nervous system is not immediately updated on changes in the peripheral motor system, and any recent commands issued may be yet to affect the musculature. Internal models are employed more rapidly, providing information about future properties of the controlled object, a fortiori in cases where accurate sensory feedback may be totally absent. This internal model 'feedback' allows the perceiver to rapidly interpret the perceptual signal and react accordingly, complete percepts received incompletely and/or under noise, and disambiguate in situations of uncertainty (Jordan & Wolpert, 2000). While internal models are conceivably located in all brain regions with synaptic plasticity that receive and send relevant information for their input and output (Kawato, 1999), a broad range of electrophysiological (Gilbert & Thach, 1977; Medina & Lisberger, 2009), imaging (Imamizu, Miyauchi, Tamada, Sasaki, Takino, Pütz, Yoshioka, & Kawato, 2000; Puttemans, Wenderoth, & Swinnen, 2005; Blakemore, Frith, & Wolpert, 2001; Desmurget, Gréa, Grethe, Prablanc, Alexander, & Grafton, 2001), and clinical studies (Müller & Dichgans, 1994; Nowak, Timmann, & Hermsdorfer, 2007) have established the cerebellum as 'the most likely site for forward models to be stored' (Kawato, Kuroda, Imamizu, Nakano, Miyauchi, & Yoshioka, 2003, p. 171).

However, the cerebellum exhibits a combination of two striking properties that support the involvement of its internal models beyond motor control: namely, its 'essentially uniform, monotonously repetitive architecture' (Schmahmann, 2000, p. 206) and its massive connectivity with virtually all major subdivisions of the brain. In particular, axonal fiber-tracing studies have identified projections from a broad range of neocortical areas to the ponto-cerebellar system, and even further to specific cerebellar lobules. These lobules return projections to the very same cerebral cortical areas via the cerebellar nuclei and thalamus, hence operating by means of segregated anatomical loops. Characteristically, the primary motor cortex selectively communicates with cerebellar lobules HV, HVI, HVIIb, and HVIIIa (Middleton & Strick, 2000; Kelly & Strick, 2003). Importantly, though, a substantial range of prefrontal areas send inputs to the pontine nuclei, which, along with the inferior olive, form the two major sources of input to the cerebellum. These prefrontal areas span from area 10 through to posterior regions of area 8, crucially including area 45B in the rostral bank of the inferior limb of the arcuate sulcus (Schmahmann & Pandya, 1997). Prefrontal neurons provide inputs to Purkinje cells in lobule VIIa and Crura I and II in lobule HVIIa (Kelly & Strick, 2003). This has encouraged the idea that '[i]f closed-loop circuits reflect a general rule, then all of the areas of cerebral cortex that project to the cerebellum are the target of cerebellar output' (Strick, Dum, & Fiez, 2009, p. 422). Furthermore, lobule HVIIa boasts unparalleled computational power: it comprises nearly 50% of the cerebellar volume (Balsters et al., 2010), while the human cerebellum itself contains more neurons than all the rest of the nervous system (see Leiner (2010) for references). Moreover, recent work employing intrinsic functional connectivity in humans has demonstrated that HVIIa Crus I/II can be further subdivided on the basis of functional connectivity with anterior, ventromedial, and dorsolateral regions in the prefrontal cortex (Krienen & Buckner, 2009; Buckner, Krienen, Castellanos, Diaz, & Yeo, 2011). Since the uniformity of cellular organization across the cerebellar cortex implies identity in the computations performed (e.g. Bloedel, 1992; Schmahmann, 1997), the same forms of plasticity might support the automation of cognitive processes, and knowledge on cerebellar motor learning has indeed started to guide the study of cerebellar contributions to cognition (Ramnani, 2006; figure 1 below).

Associative learning represents a fundamental mechanism by which cerebellar internal models are seen to work. Classical conditioning, for instance, provides the most basic form of associative memory formation, and the cerebellum has long been established as a fundamental site (e.g. Thompson, Bao, Chen, Cipriano, Grethe, Kim, Thompson, Tracy, Weninger, & Krupa, 1997; Christian & Thompson, 2005) in both comparative (e.g. McCormick & Thompson, 1984) and human clinical studies (e.g. Daum, Ackermann, Schugens, Reimold, Dichgans, & Birbaumer, 1993; Timmann, Kolb, Baier, Rijntjes, Mueller, Diener, & Weiller, 1996). Crucially, the cerebellum is involved in non-motor aspects of associative learning. In a series of studies, cerebellar patients acquired associations between colors and numerals by trial-and-error. In comparison with normal controls, patients were significantly slower in learning the correct associations, and were impaired in recognizing them later. Control conditions established that this learning deficit could not be reduced to patients' motor impairment (Drepper, Timmann, Kolb, & Diener, 1999; Timmann, Drepper, Maschke, Kolb, Böring, Thilmann, & Diener, 2002; Timmann, Drepper, Calabrese, Bürgerhoff, Maschke, Kolb, Daum, & Diener, 2004). Moreover, recent fMRI work has shown that HVIIa Crura I/II are involved in the acquisition and employment of first- and second-order rules. In these studies, subjects acquired arbitrary associations by trial-and-error of visual stimuli with manual responses (first-order rules) or with instructions on the selection of a first-order rule (second-order rules). With the employment of delayed response tasks, activity time-locked to the onset of the rule-related cue was disambiguated from that for motor responses or visual feedback (Balsters & Ramnani, 2008, 2011; Balsters, Whelan, Robertson, & Ramnani, 2013). These findings corroborate early proposals that the cerebellum 'predicts' and 'prepares' the internal conditions required for sensory, motor, autonomic, memory-related, attention-related, affective, or linguistic operations, by acquiring the 'predictive relationships among temporally ordered multidimensional sequences of exogenously derived […] and endogenously derived […] neural activities' (Courchesne & Allen, 1997, p. 2). More recently, contributions of cerebellar internal models to multi-modal associative learning have been thoroughly examined in two reviews (Timmann, Drepper, Frings, Maschke, Richter, Gerwig, & Kolb, 2010; Bellebaum & Daum, 2011). Quite importantly, though, neither of those discusses the ways in which cerebellar internal models could contribute to non-motor aspects of linguistic associative learning and processing.

[Please insert figure 1 here]

3. 'Cerebellum and Language', but no 'Prediction'

In parallel, studies on the cerebellar contributions in language have started to flourish over the last few decades. Although still often marginalized in neurobiological models of language processing (e.g. Hickok & Poeppel, 2007; Friederici, 2012), the 'ongoing enigma' of the linguistic cerebellum has been attracting significant attention (Mariën et al., 2014). Indeed, there has been cumulating clinical evidence for even a ‘lateralized linguistic cerebellum’ (Mariën, Engelborghs, Fabbro, & De Deyn, 2001), and meta-analyses of imaging studies locate activations related to language processing in right hemispheric cerebellar regions (E, Chen, Ho, & Desmond, 2014; Stoodley & Schmahmann, 2009). For an up-to-date account of the evidence, the reader is encouraged to consult De Smet, Paquier, Verhoeven, and Mariën (2013), as well as papers in this special issue. Even so, such discussion has not incorporated the way in which cerebellar internal models may be involved in language processing, until very recently.

4. ‘Language and Prediction’, but no ‘Cerebellum’

Turning to prediction mechanisms in language processing, these are anything but novel a concept in studies of psycholinguistics and the neurobiology of language. For decades now, research has shown that predictions at the sentence level modulate speech perception and production (Lieberman, 1963) and accelerate syntactic processing (e.g. Staub & Clifton, 2006; see also discussion in Sturt & Lombardo, 2005).

Unsurprisingly, then, prediction in language processing is often conceptualized to operate in the form of internal model outputs. These are used in work on auditory feedback for speech production (Perkell, Matthies, Lane, Guenther, Wilhelms-Tricarico, Wozniak, & Guiod, 1997; Golfinopoulos, Tourville, Bohland, Ghosh, Nieto-Castanon, & Guenther, 2011; Golfinopoulos, Tourville, & Guenther, 2010), providing rapid information on the predicted auditory consequences of articulatory gestures well before the later-arriving sensory feedback (Heinks-Maldonado, Nagarajan, & Houde, 2006; Tourville, Reilly, & Guenther, 2008), or even in the absence of overt articulation (Tian & Poeppel, 2010). The N400, an event-related brain potential response, has also been discussed within the context of violated predictions generated by internal models (Hosemann, Herrmann, Steinbach, Bornkessel-Schlesewsky, & Schlesewsky, 2013).

In their seminal paper, Pickering and Garrod (2007) advanced the idea that internal models are employed in language processing in a multi-level fashion, extending beyond acoustic perception. Language perception and comprehension were proposed to dynamically combine the outputs generated by an ‘input analysis system’ with predictions of internal models providing an estimate of the next state that the input analysis system will enter (figure 2 below). Whenever the prediction is strong and the input noisy, the internal model exerts strong influence to the ‘input analysis system’. In that way, the internal model ensures noise-resistant, rapid perception and comprehension of utterances. It should be noted that Pickering and Garrod’s (2007) work draws from literature discussing the covert, imitative involvement of action production mechanisms in the efficient perception of actions performed by conspecifics (Wilson & Knoblich, 2005). Whether cerebellar internal models in language comprehension recruit language generation mechanisms remains an outstanding question (see section 7 below).

[Please insert figure 2 here]

5. Cerebellum, Prediction and Language

Given the well-established nature of cerebellar contributions in motor control (Manto et al., 2012), studies of speech production and perception have synthesized insight from these three strands of research (cerebellar involvement in predictive processing; engagement of the cerebellum in ‘non-motor’ aspects of language; prediction mechanisms in language processing). Recent proposals emphasize the importance of the cerebellum and the basal ganglia in detecting temporal regularities and generating predictions during speech processing- hence alleviating the workload of fronto-temporal speech processing networks (Kotz & Schwartze, 2010). Accounts of language production also suggest that cerebellar internal models operate at the articulatory level (Hickok, 2012). Superior paravermal regions, for instance, may encode feedforward programs for the production of syllables (Ghosh, Tourville, & Guenther, 2009). Stuttering has been discussed within the context of deficient cerebellar internal models in generating a motor prediction error (Golfinopoulos, Tourville, & Guenther, 2010). Similarly, right cerebellar lobule HVI may undertake the temporal processing of segmental properties and left HVI may process prosodic, melodic properties (Callan et al., 2007). More recent work implicates the cerebellum in perceptual enhancement of acoustic properties of the linguistic signal. An fMRI study employed a word recognition task including acoustic stimuli from severely distorted speech. Improvements in the perception of such stimuli modulated cerebellar activity in four distinct cerebellar regions, one crucially being the right HVIIa Crus I. Activation in this lobule functionally correlated with cerebral regions that encompassed portions of the left angular and left temporal gyri. The findings supported the idea that regions within the left temporal and parietal cortex and the right Crus I (potentially along with lobules V/VI), participate in a functional network for achieving adaptive plasticity in speech perception. Discrepancies between the actual distorted acoustic speech input and the predicted acoustic input for a lexical item were discussed as engaging cerebellar-dependent supervised learning mechanisms (Guediche, Holt, Laurent, Lim, & Fiez, 2014).

The evidence accumulated so far has already encouraged discussions on the extension of the contributions of cerebellar internal models to non-motor aspects of language processing (Ito, 2000a, 2008; Argyropoulos, 2008, 2009). For instance, Argyropoulos (2009) argued that the cerebellum can store associative memory traces of contiguous linguistic events in sentence comprehension, such as the instance of processing the phonological, syntactic, and semantic properties of word ‘x’ and subsequently those of word ‘y’. In brief, cerebellar internal models would receive a copy of the input processed at each stage by a fronto-temporal ‘input analysis system’ (Pickering & Garrod, 2007) along with the operations performed on it, via the cerebro-ponto-cerebellar pathway. The internal models would then output a prediction of the next stage that the system would enter, via the ventrolateral neodentate, the parvocellular red nucleus and the thalamus back to the prefrontal cortex. Any discrepancies between the two would be conveyed as error signals back to the neocerebellar cortex through climbing fibers from the inferior olive. By long-term depression of parallel fiber-to-Purkinje cell synapses, these errors would train the internal model for accurate predictions. In cases where input analysis is conducted in noisy conditions and/or strong neocerebellar predictions are transmitted, neocerebellar output would override and/or bypass the output of cortico-cortical processing. With its massive computational power, lobule HVIIa may successfully undertake such an enormous task. Figure 3 provides a schematic illustration of phonological predictions generated in cerebro-cerebellar circuitry-conceivably, though, cerebellar circuitry would undertake the acquisition of temporal regularities and the implementation of predictions at different levels (e.g. phonology, syntax, semantics). While the specialization of distinct lobular regions in different levels of prediction remains an outstanding question in the field (see section 7 below), a series of studies have followed these proposals, providing evidence for the possibility of cerebellar internal models to transmit predictions on upcoming ‘non-motor’ information in sentence processing.

[Please insert figure 3 here]

5.1. Cerebellar involvement in phrasal predictions

In the first cerebellar TMS study to use a language task, Argyropoulos (2011a) addressed cerebellar contributions to such ‘non-motor’ aspects of predictive language processing. At a minimal level, the capacity of the neocerebellum to bias and/or pre-empt the prefrontal cortex for upcoming linguistic input was predicted to manifest itself in phrasal associative, and not semantic categorical lexical priming. On the one hand, phrasal associative priming pertains to the probability that one word may call to mind a second one, and is based on the temporal contiguity and predictability of items in discourse. Pairs of lexical morphemes in idiomatic phrases provide a case in point, such as ‘gift-horse’, ‘skeletons-closet’. On the other hand, semantic categorical relatedness reflects the taxonomic relations between paradigmatic co-exemplars and the overlap in featural descriptions of two words, e.g. ‘tree’-‘bush’, ‘bee’-‘grasshopper’ (Thompson-Schill, Kurtz, & Gabrieli, 1998; Hutchison, 2003).

A lexical decision task assessed noun-to-noun priming sizes (quantified as differences in milliseconds between decision latencies for related and unrelated lexical pairs) before and after rTMS of the right neocerebellar vermis and of a control site in healthy native English speakers, as yielded by two different types of pairs: phrasal associates (e.g. ‘gift’-‘horse’), and pairs of subordinate and superordinate terms of the same category (e.g. ‘penny’-‘coin’). TMS of the neocerebellar vermis selectively enhanced phrasal associative priming in the form of increased differences between response latencies for lexical targets preceded by an unrelated prime (e.g. ‘battery’-‘horse’) and those preceded by a phrasally associated one (e.g. ‘gift-‘horse’). No such effects occurred for categorically related items after TMS, and stimulation of the control site did not affect associative priming sizes. This finding was consistent with evidence that cerebellar damage impairs verbal fluency by affecting phonemic rule-based word production, yet sparing semantic rule-based performance (Leggio, Silveri, Petrosini, & Molinari, 2000). Encouragingly, research in Alzheimer's disease and semantic dementia has established that temporal lobe lesions induce disruptions in categorical semantic, but not in associative priming (Rogers & Friedman, 2008). Moreover, cerebellar patients have been shown to ably perform category learning tasks, unlike patients with Parkinson's disease (Maddox, Aparicio, Marchant, & Ivry, 2005). This finding also echoes speculations on the significance of the cerebellum in storing and generating prefabricated, rote-memorized, idiomatic sequences, as opposed to propositionally composed ones: '[w]e learn and can recite ''Jabberwocky'' as movement and not at all as language […] We rote- memorize something that has so little linguistic or logical connection among the elements that it is learned as a movement. We can listen to what we say in order to get at what we otherwise can't remember […] But it is not something we know. It is buried in a rote-learning movement sequence' (Thach, 1997, pp. 91-2).

5.2. Cerebellar involvement in semantic predictions

Arguably, though, word-to-word predictions need not occur exclusively at the phonological level. There is substantial evidence in the lexical priming literature for semantic associations reflecting the predictive linkage of concepts based on world knowledge, as in 'instrument-action' pairs ('broom'-'sweep'), 'script relations' ('theatre'-'play'), 'locative relations' ('beach'-'house'), 'compositional relations' ('brick'-'house'; Hutchison, 2003; Jones & Estes, 2012). According to the theory on event schemata, stereotypical knowledge about the world is organized into units, including information about sequences of objects, situations, events, or states. Schemata are not the products of inferential manipulation of declaratively encoded representations. They operate online and below the level of awareness in sentence comprehension upon encountering the relevant concepts (e.g. the association of a dog with the agent-role in a biting event). In both dual-stage and constraint-based models of sentence comprehension, mechanisms of prediction (McRae, Hare, Elman, & Ferretti, 2005) and schema transmission (Ferreira, 2003; Townsend & Bever, 2001) employ these semantic associations in order to assign probable thematic (θ)-roles (Chomsky, 1981) and form a meaning-form hypothesis.

Suggestively, there is ample evidence from imaging, neurostimulation, and clinical studies in support of the involvement of the cerebellum in word generation tasks, crucially verb-to-noun generation. In a seminal PET study, subjects read aloud, repeated, passively read, or listened to nouns, or generated semantically appropriate verbs in response (e.g. ‘eat’ for ‘cake’). Blood flow changes occurred in paravermal regions when subjects read aloud or repeated nouns, as compared to when they viewed or listened to them. However, activation in the right lateral cerebellum was found for overt verb generation, but not for noun repetition/reading. This was difficult to account for on a motor basis. It was equally enigmatic though with respect to the involvement of the cerebellum in computations underlying verb generation (Petersen, Fox, Posner, Mintun, & Raichle, 1989). Recent fMRI studies provide similar results. Comparing verb generation and verb reading, Frings, Dimitrova, Schorn, Elles, Hein-Kropp, Gizewski, Diener and Timmann (2006) found activations in the right cerebellar lobule HVI and HVIIa Crus I as a measure of verb generation; again, these were lateral from the paravermal activation of lobule VI, which was associated with speech articulation. Fiez, Petersen, Cheney, and Raichle (1992) examined an English-speaking patient with a large right cerebellar infarct, who generated inappropriate responses in a number of word-generation tasks, involving verb generation. The patient’s responses were inappropriate, but remained categorically related (e.g. ‘small’, instead of take or ‘swallow’, in response to ‘pill’). This could not be attributed to overall cognitive impairment, as the patient’s performance on tests of memory, intelligence, ‘frontal function’, and language skills was excellent. This supports the idea that cerebellar damage leaves semantic networks intact (Fiez & Raichle, 1997). In another study, patients performed poorly in generating appropriate verbs for nouns, but selected the correct verb for a noun from a list of alternative responses, suggesting that semantic/syntactic representations were preserved. They were also able to produce appropriate subordinate term-responses to superordinate terms, suggesting that ‘[t]he right posterolateral cerebellum may be more involved in associative semantics than in categorical semantics’ (Gebhart, Petersen, & Thach, 2002). This pattern of impairments is qualitatively different from the direct disruptions in semantic networks associated with temporal cortical pathology in semantic dementia and Alzheimer’s disease (e.g. Rogers & Friedman, 2008). Studies on ‘associative’ and ‘semantic’ priming in Broca’s and Wernicke’s aphasia also show no impairments in automatically accessing lexical and/or semantic information (Blumstein, Milberg, & Shrier, 1982; Milberg & Blumstein, 1981), suggesting that associative priming does not rely on these perisylvian structures. As already discussed, cerebellar patients ably perform category learning tasks, unlike patients with Parkinson's disease (Maddox et al., 2005). Moreover, in a study employing a card-sequencing task, cerebellar patients with left lesions showed selective impairments in script sequences based on pictorial material, while those with right lesions were only impaired in script sequences requiring verbal elaboration (Leggio, Tedesco, Chiricozzi, Clausi, Orsini, & Molinari, 2008). Likewise, cerebellar cathodal tDCS has facilitated the rate and consistency of responses in a verb generation task, as compared to sham and anodal stimulation. These facilitatory effects occur in the case of cerebral cortical tDCS following anodal stimulation. They were hence explained in terms of disinhibition of the left prefrontal cerebral cortex resulting from the inhibitory effect of cathodal tDCS on the overall inhibitory tone that the cerebellum exerts on the cerebral cortex (Pope & Miall, 2012). Finally, in a TMS study, right cerebellar rTMS reduced category switching in the form of reduced phonemic and semantic fluency (Arasanz, Staines, Roy, & Schweizer, 2012).

The involvement of neocerebellar circuitry in the generation of semantic predictions was first directly assessed in a cerebellar rTMS study (Argyropoulos, 2011b; Argyropoulos & Muggleton, 2013). A pair-wise lexical decision task used noun-primes, the semantic properties of which could be categorically related by synonymy (e.g. ‘theft’-‘stealing’) or thematically associated (denoting agents, patients, instruments or locations of actions) with their verb-targets (e.g. ‘chef’-‘cooking’). Four groups of subjects were employed: a group that underwent stimulation of a medial cerebellar site; a group that underwent stimulation of a right lateral cerebellar site corresponding to a region in right HVIIa Crus I involved in cognitive aspects of verb generation (Frings et al., 2006); two groups that completed the same session with no intervention of TMS, one with a 5-minute break and another one without. Stimulation of this lateral cerebellar site selectively boosted associative priming. No effects appeared after medial cerebellar stimulation or no stimulation. Argyropoulos (2011b) and Argyropoulos and Muggleton (2013) discussed this effect as a disruption of inhibitory processes in generating predictions for associated events expressed by the verb-target, e.g. (cooking)´, upon processing the semantic properties of the noun-prime, e.g. (chef)´.

This study was soon followed up by Lesage, Morgan, Olson, Meyer, and Miall (2012; see also Lesage (2013) for stimuli used). The authors employed the ‘Visual World’ paradigm (Altmann & Kamide, 1999), assessing the onset latencies of listeners’ saccadic eye movements towards images of objects that aurally presented sentences referenced. The authors used two types of sentences: ‘predictive’ sentences, where the object could be predicted among four alternatives (one target object and three distractor objects) on the basis of the semantic content of the verb (e.g. ‘The man will sail the … boat/mountain/bird/car’), and control, ‘non-predictive’ sentences (e.g. ‘The man will watch the … boat/mountain/bird/car’), where the verb was not selective for the target object. In pre-TMS conditions, priming occurred in the form of faster anticipatory responses in the predictive as compared to the non-predictive condition. Cerebellar rTMS significantly delayed such anticipatory responses. Stimulation did not change saccadic latencies in the non-predictive sentences, ruling out a general effect on language processing. It did not change eye movement kinematics either, thus ruling out disruptions in oculomotor control. The prediction deficit was moreover absent in two control groups (vertex stimulation and no stimulation), ruling out non-specific effects of stimulation.

More recently, two fMRI studies have attempted to address this issue. Both used sentence comprehension tasks manipulating cloze probabilities and hence the strength of ‘semantic priming’ (Moberget, Gullesen, Andersson, Ivry, & Endestad, 2014; Lesage, Hansen, & Miall, 2014; unpublished). The task of the first study involved the presentation of a sequence of five centrally presented words. The authors manipulated cloze probability by varying the context provided by the sequence of the initial four words. In a ‘Congruent’ condition, the target word was highly predictable on the grounds of the preceding four (e.g., ‘two plus two is four’). In an ‘Incongruent’ condition, a final word was also highly predictable, but that prediction was violated by the target word presented (e.g., ‘[the water] had frozen to cars’). In a ‘Scrambled’ condition, the first four words made a non-grammatical sentence stem (e.g., ‘fast in clock plane’), rendering the target word unpredictable (e.g., ‘through’). In the end of each presentation, participants assessed whether the sentence was meaningful or not by a button press. In their contrasts of interest, ‘Congruent > Scrambled’ yielded a cluster of activation across right HVIIa Crus I/II.

‘Incongruent > Congruent’ yielded an activation cluster across left HVIIa Crus I/II, IV, and medial VIIa Crus I and another across right HVIIa Crus I/II, medial VIIa Crus I/II and VI. ‘Incongruent > Scrambled’ showed activation in right HVIIa Crus I/II, VIIa Crus I/II, and VI, and left HVIIa Crus I/II and VIIa Crus I. The authors reasoned that both the generation of a prediction for an upcoming word based on its sentential context as well as the violation of that prediction are consistent with the involvement of neocerebellar internal models in non-motor aspects of language processing (Moberget et al., 2014).

Using an event-related fMRI design, the authors of the second study demonstrated that sentence predictability modulated the BOLD signal amplitude in a right HVIIa. Three events were modelled per trial: a context sentence (e.g. ‘Greg went home for Christmas dinner’); the stem of a second sentence (e.g. ‘His mum always cooked a’); the end of the second sentence (e.g. ‘turkey’). Subjects read the sentences and pressed a button indicating the plausibility of the outcome. Crucially, the stem event did not require a motor response. Predictability was used as a parametric modulator for context and stem events. The authors showed that ‘a cluster in the right posterolateral cerebellum […] was modulated by the predictability of the stem independent of outcome’ (Lesage, Hansen, & Miall, 2014; see Lesage (2012, pp. 194-5) for stimuli).

5.3. *Cerebellar contributions to the routinization of language processing*

Cerebellar internal models have been held not only to generate predictions about the next state of a simulated process, but also to gradually undertake its automatic implementation (Ito, 1984; 2008). Interestingly, casual adult dialogical interaction exhibits phenomena of routinized language processing, with lexical items and syntactic constructions of re-occurring phonological and semantic properties (Pickering & Garrod, 2004). The significance of the cerebellum in automatizing motor repertoires has been supported by clinical (e.g. Lang & Bastian, 2002) and imaging evidence (e.g. Jueptner, Frith, Brooks, Frackowiak, & Passingham, 1997; Toni, Krams, Turner, & Passingham, 1998), extending recently to higher cognitive aspects of rule learning (e.g. Balsters & Ramnani, 2011). In a tDCS study, cerebellar stimulation impaired the practice-dependent improvement in performance in the Sternberg task irrespective of polarity. On the contrary, tDCS over the dorsolateral prefrontal cortex left it unchanged, markedly affecting verbal working memory per se.

Cerebellar stimulation also left visual evoked potentials unchanged, thus excluding visual cortex involvement (Ferrucci, Marceglia, Vergari, Cogiamanian, Mrakic-Sposta, Mameli, Zago, Barbieri, & Priori, 2008).

Crucially, cerebellar impairments induce similar automatization deficits in language-related processes. In word generation tasks, cerebellar patients reduce their response latencies poorly across repeated blocks of trials in comparison with normal controls (Fiez et al., 1992; Gebhart, Petersen, & Thach, 2002). Further evidence has been recently provided by Argyropoulos, Kimiskidis, and Papagiannopoulos (2011). The authors recruited two groups of participants of native speakers of Modern Greek for a TMS study conducted in two sessions, with the same stimuli presented in the second session. One group received stimulation centered at a site overlying the neocerebellar vermis in the first session, and on a deeper, control site in the second. Another group received stimulation in these two sites in reverse order. The study employed a pairwise lexical decision task. Lexical decision latencies for the group that underwent stimulation of the neocerebellar vermis in the second session did not become any shorter after stimulation, in contrast to all other conditions. Such disruption could not be explained on the grounds of sensorimotor processes, since the second group, who received neocerebellar vermal stimulation in their first session, showed a significant reduction of their reaction times in that session after stimulation. However, the disruption only appeared after stimulation of the target site selectively for the first group, who were encountering letter-string pairs for the second time. The automatization deficit thus pertained in particular to processing pairs that had been already encountered.

6. *Critical assessment*

The evidence presented here for the involvement of the cerebellum in ‘non-motor’ aspects of predictive language processing is consistent with the unitary, multi-modal nature of the contributions of cerebellar internal models. It appears also to follow from the sizeable volume of suggestive evidence that has preceded it. However, as this section will demonstrate, there are significant issues that compromise the strength of these findings and limit their interpretation.

6.1. *TMS studies*

Despite the significant advantages of cerebellar neurostimulation with TMS (Grimaldi et al., 2013) and tDCS (Grimaldi et al., 2014) over cerebellar patient and fMRI studies, there are methodological issues that are often hard to overcome. For instance, in TMS, optimal localization of the target region is meaningfully achieved by retrieving coordinates of the area with the highest level of activation during performance of the same task in a previously completed fMRI session by the very same subject. Alternatively, high-resolution structural images may be used to position the coil according to the subject's underlying anatomy. When these are unavailable, scalp-based measurements from clear external landmarks are often used, such as the vertex or the inion (Devlin & Watkins, 2007). Indeed, all four TMS studies discussed here (Argyropoulos, 2011a; Argyropoulos, Kimiskidis, & Papagiannopoulos, 2011; Argyropoulos & Muggleton, 2013; Lesage et al., 2012), used the inion as an external landmark, thus making precise localization of the lobular cortex stimulated impossible. However, all four studies demonstrate further weaknesses that interact with the aforementioned limitation and compromise the inferences that can be drawn.

6.1.1. Stimulating the neocerebellar vermis

The putative target in the study of Argyropoulos (2011a) and Argyropoulos, Kimiskidis, and Papagiannopoulos (2011) was a right superior posterior vermal site. Their external landmark coordinates (1 cm below the inion, 1 cm laterally to right) targeted the medial (VIIa) Crus I/II. In order to estimate the depth of the site and thus the possibility of its successful stimulation, the authors recruited a volunteer whose brain image was already registered with a TMS-MRI co-registration system. The site corresponded to the right VIIa. Encouragingly, the superior posterior vermis is one of the lobules closest to the TMS coil (Miall & Christensen, 2004), and its stimulation has induced behavioral effects with high spatial precision (Hashimoto & Ohtsuka, 1995).

However, the anatomical connectivity of the vermis with the frontal lobes remains unclear (e.g. Kelly & Strick (2003) on medial VIIa-BA 46 connectivity; but also Coffman, Dum, & Strick (2011) on vermal-motor cortical connectivity). On the other hand, resting-state functional connectivity studies demonstrate that the vermis and the medial regions of the Crura contain a broad range of network nodes, the most posteromedial of which belongs to the dorsal attention network (Buckner et al., 2011). In another study, a cluster spanning across Crus II with the medial portions close to the vermis showed connectivity with the left executive control network (Habas, Kamdar, Nguyen, Prater, Beckmann, Menon, & Greicius, 2009). More recently, Halko, Farzan, Eldaief, Schmahmann, and Pascual-Leone (2014) applied cerebellar rTMS guided by subject-specific connectivity to evaluate the relevance of connections between cerebral and cerebellar hubs belonging to different functional networks. One of their stimulation sites was lobule VII, and they demonstrated that its stimulation influences the cerebral dorsal attention system.

Equally enigmatic is the nature of vermal contributions in language. On the one hand, vermal lobule VII is involved in oculomotor control (e.g. Hashimoto & Ohtsuka, 1995) and also supports emotional processing (E, Chen, Ho, & Desmond, 2014; Stoodley & Schmahmann, 2009). On the other hand, superior posterior vermal activations are found in studies of language processing that are hard to reduce to motor effects (Desmond & Fiez, 1998; Desmond, Gabrieli, & Glover, 1998). Vermal tumor resection in children is associated with agrammatism (Riva & Giorgi, 2000). Voxel-based morphometric studies also report correlations between vermal grey matter and working memory measures (Ding, Qin, Jiang, Zhang, & Yu, 2012), and other studies find increased activity in the vermis after lexical training (Raboyeau, Marie, Balduyck, Gros, Démonet, & Cardebat, 2004). Vermal atrophy is also the most widely-cited cerebellar abnormality in schizophrenia (Picard, Amado, Mouchet-Mages, Olié, & Krebs, 2008): indeed, lexical priming studies on schizophrenic patients show greater priming for associated-only word pairs than for pairs related only semantically or both semantically and associatively (Nestor, Valdman, Niznikiewicz, Spencer, McCarley, & Shenton, 2006).

Evidently, though, the use of external landmarks for coil positioning could not exclude the probability that adjacent lobular cortices of VI, VIIb and VIII were also stimulated. While stimulation of right paravermal compartments is traditionally achieved by placing the coil 2 cm laterally to the right from the inion (e.g. Miall & Christensen, 2004) instead of 1cm, the scalp coordinates used here for targeting the right neocerebellar vermis may concomitantly stimulate portions of the right superior paravermis (e.g. Hashimoto & Ohtsuka, 1995). These lobules are reciprocally connected with the motor cortex in a segregated fashion (Kelly & Strick, 2003). While the second priming type used (semantically related and unrelated pairs) provided sufficient conditions for controlling for motor effects, the interpretation of the main effect here would be different. It can only be speculated that the selective effect on associative priming would pertain to cerebellar involvement in ‘silent/covert speech’ (e.g. Ackermann, Wildgruber, & Grodd, 1998). In an fMRI study, activation within the same region was found during silent recitation of the names of the months of the year (Ackermann, Mathiak, & Ivry, 2004), i.e. a routinized sequence of lexical morphemes. Such ‘inner speech’ has been held to provide a common platform for a broad range of cognitive functions implicating the cerebellum (Ackermann, Mathiak, & Riecker, 2007). This account would echo discussions whereby formulaic utterances are ‘buried in a rote-learning movement sequence’ engaging the cerebellum (Thach, 1997, pp. 91-2). In that way, it could only be speculated that the abnormalities in predictive functions were owed to impairments in the covert employment of the language production circuit (Pickering & Garrod, 2007).

For the study in Argyropoulos (2011a), in particular, a series of other weaknesses would include the low number of subjects and items per condition. Moreover, the within-subjects design involved participants being exposed to the same items twice (stimulation of control and target site counterbalanced across subjects). This may have introduced confounds with the effects that cerebellar stimulation has on practice-induced facilitation in performance (Ferrucci et al., 2008; Argyropoulos, Kimiskidis, & Papagiannopoulos, 2011). Also, the study in Argyropoulos, Kimiskidis and Papagiannopoulos (2011) failed to replicate the selective effects of neocerebellar vermal stimulation on associative priming. This could be attributed to the unavailability of published word association norms in Modern Greek for the construction of the stimulus set. On the other hand, the main finding in Argyropoulos, Kimiskidis, and Papagiannopoulos (2011) on the effects of neocerebellar vermal stimulation on the disruption of practice-induced accelerations of lexical decisions when participants encountered the same stimuli for the second time was not present in Argyropoulos (2011a). This could be due to the higher number of participants employed in that study.

6.1.2. Stimulating the neocerebellar hemispheres

Similarly, in Argyropoulos & Muggleton (2013), the target stimulation site was identified by external landmarks. The aim was to stimulate the peak coordinates of the activation cluster yielded in Frings et al. (2006) when comparing verb generation in inner speech with verb reading in inner speech (right HVIIa Crus I). Its coordinates were again registered and converted into scalp coordinates using a TMS-MRI co-registration system in one volunteer, and corresponded to 10 cm laterally to the right from the inion. The significant distance from that landmark would only increase the inter-subjective variability and the uncertainty on the cerebellar lobule stimulated. Moreover, the significant depth of the underlying tissue may have made it impossible to successfully induce an effect in at least a subset of participants with the figure-of-eight coil used (Hardwick, Lesage, & Miall, 2014), despite the high amplitude of stimulation. Another concern is that significant associative priming was only observed in the second phase of the experimental session across the four groups (control TMS, target TMS, no TMS 1, no TMS 2). While the associative boost after stimulation of this lateral site was significantly larger than that for the other groups and was also absent from the semantic categorical priming type, a concrete interpretation of the effect remains elusive. It would suggest that the processes affected were not automatic in nature and may reflect TMS effects on strategically employing semantic associations.

Similar issues in identifying the exact cerebellar lobular cortex underlying the stimulation site apply in Lesage et al*.* (2012), given that external landmarks were used again. However, the authors successfully stimulated the deeper hemispheric regions of the neocerebellum by employing a double-cone coil (Hardwick, Lesage, & Miall, 2014). As compared to figure-of-eight coils, though, focality of stimulation is more limited, and the possibility of a spread of the effect in neighbouring areas is increased. Suggestively, the authors used a set of coordinates (1 cm below the inion and 3 cm laterally to the right) that have been traditionally employed to induce cerebellar inhibition and in return facilitation in MEPs evoked by single pulse TMS of the contralateral primary motor cortex (e.g. Oliveri, Koch, Torriero, & Caltagirone, 2005). These coordinates have also been used by the same laboratory in recent assessments of successfully modulating motor cortical excitability with cerebellar stimulation by different coil types (Hardwick, Lesage, & Miall, 2014). They have also been employed in their previous TMS investigations to aim at the hand area of the ipsilateral cerebellar cortex, with its stimulation affecting cortico-cerebellar projections, changing motor cortical excitatability in the contralateral hand area (Miall, Christensen, Cain, & Stanley, 2007). Previous studies have demonstrated that this target site ‘affects the hand area of motor cortex, consistent with activation of lateral cerebellar cortex, probably in lobules V and VI’ (Miall & King, 2007, p. 576). Therefore, despite the fact that in Lesage et al*.* (2012) stimulation with the same coordinates was ‘directed towards Crus II’ (Grimaldi et al., 2013, p. 133), motor-projecting cerebellar lobules HV and HVI were clearly also stimulated.

However, this limitation further interacts with a weakness in the construction of the stimulus set of the task. The authors acknowledge that they did not distinguish action-related from non-action related verbs, and were thus unable to test for any specific effects of action verbs in cerebellar motor-projecting lobules. They entertain the possibility that these are more active when processing action-related verbs. They reason, however, that, if such mechanisms were indeed present, they would not suffice to explain their results, given that the impairment was specific to sentences with predictive verbs, irrespective of any action-related semantics. The certainty with which such an interpretation can be dismissed comes into question when the stimuli (Lesage, 2013, p. 189) used in this study (Lesage et al., 2012) are examined under further scrutiny. The verbs in the ‘predictive’ condition score higher in concreteness, imageability, and meaningfulness metrics (e.g. Coltheart, 1981; Brysbaert, Warriner, & Kuperman, 2014) than those in the ‘non-predictive’ (control) condition. This is illustrated in table 1 below. The cerebellar lobules HV and HVI that were concomitantly stimulated form a segregated loop with the motor cortex (Kelly & Strick, 2003), and ample evidence associates action-related semantic processing with activity in motor structures used to implement those actions. For instance, processing speed for leg-related words is affected by TMS over the leg area of the left primary motor cortex, while performance remains unaffected when the left arm area or the right leg area is stimulated (Pulvermüller, Hauk, Nikulin, & Ilmoniemi, 2005). Similarly, fMRI studies show activations in passive reading of verbs that denote actions executed by different effectors are somatotopically organised in the primary motor cortex (e.g. Hauk, Johnsrude, & Pulvermüller, 2004). Another study has demonstrated that the MEP size in each effector muscle is only affected when listening to sentences containing actions related to that effector (Buccino, Riggio, Melli, Binkofski, Gallese, & Rizzolatti, 2005). Moreover, cathodal tDCS on the left motor cortex in healthy participants has reduced success rates in the acquisition of action-related words, in comparison with anodal or sham stimulation, tDCS over the prefrontal cortex, and learning of object-related words (Liuzzi et al., 2010). Unsurprisingly, imagined movement also engages the cerebellum (e.g. Hanakawa et al., 2008). Some first suggestive evidence for the modulation of processing action-related verbs by cerebellar TMS can be found in Oliveri et al. (2009).

[Please inset table 1 about here]

What if stimulation was to selectively affect HVIIa Crus I/II, which communicates with the contralateral dorsolateral prefrontal cortex (Kelly & Strick, 2003)? The confounds in the design of the stimuli between sentence 'predictiveness' on the one hand and verb imageability, concreteness, and meaningfulness on the other would still not support the conclusion that cerebellar TMS disrupted cerebellar predictive processing. This is because of the modulation of activity in the dorsolateral prefrontal cortex by such lexical semantic properties. Concrete as compared to abstract concepts elicit greater activity in the left dorsolateral prefrontal cortex (Roxbury, McMahon, & Copland, 2014). Stimulation of the left dorsolateral prefrontal cortex also preferentially affects verbs compared to nouns (Shapiro, Pascual-Leone, Mottaghy, Gangitano, & Caramazza, 2001; Cappa, Sandrini, Rossini, Sosta, & Miniussi, 2002). This supports the interpretation that the functional link with motor cortical hand or leg regions may be mediated by the dorsolateral prefrontal cortex, much as the ventrolateral prefrontal cortex mediates the connection between speech perception and the motor cortical mouth region (Watkins & Paus, 2004): rTMS of the left dorsolateral prefrontal cortex has decreased naming latencies for verbs relative to its right homologue and sham stimulation, leaving latencies for object naming unaffected (Cappa et al., 2002). In another experiment, participants inflected nouns and verbs of a natural language as well as their pseudoword counterparts. Stimulation over the left dorsolateral prefrontal cortex affected reaction times selectively for verbs and not nouns (Shapiro et al., 2001). Therefore, the selective effect that cerebellar TMS had on 'predictive' sentences may be explained by the fact that the properties of the verbs in these sentences engaged to a significantly larger extent action-related semantic processing in the motor and/or dorsolateral prefrontal cortex than the more abstract and less imageable and meaningful verbs in 'non-predictive' sentences did. Stimulation of motor- or prefrontal-projecting cerebellar lobules may have thus disrupted this engagement.

6.2. fMRI studies

Turning to the two fMRI studies recently reported (Moberget et al., 2014; Lesage et al., 2014), a concern in the design of both is the confound between processing difficulty and ‘predictability’. As it shall be argued, the activations in the particular regions of HVIIa Crus I/II that the authors report are probably yielded by differences in generic processing demands and not predictability per se.

6.2.1. Events: predictive or predicted?

A question for both fMRI studies would pertain to the nature of the events of interest. The activations in both studies relate to activity time-locked to the onset of a stimulus (word or sentence stem) that is constrained to a smaller or larger extent by the preceding context. It is therefore not clear whether those activations reflect cerebellar involvement in generating predictions for upcoming language input or cerebellar modulation by the processing demands for input constrained to a smaller or larger extent by preceding context.

Moreover, in Moberget et al. (2014) the activity at the onset of the final word was not sufficiently disambiguated from and was susceptible to contamination by that for response preparation and implementation. Their sentence types did not only differ with respect to the predictability of the last word, but also with respect to the demands in response preparation. The activations they report in the contrast ‘Incongruent > Congruent’ provide a characteristic case in point. Indeed, the activations in this contrast fall within the region of lobule HVIIa Crus I/II that shows functional connectivity with Krienen and Buckner’s (2009) dorsolateral prefrontal cortical map and the frontoparietal control network (Buckner et al., 2011). However, this contrast need not reflect a violation of prediction in sentence comprehension. The difference in predictability is confounded with a number of discrepancies in other demands. Incongruent sentences here invite increased attention to and inhibitory control of an action being prepared (button key press to signal positive response to the question of meaningfulness) but needs to be revised (different button key press to signal negative response) upon encountering the last word; elevated demands for checking morpho-syntactic feature consistency and achieving semantic integration; elevated verbal working memory demands in reanalysing the semantic and syntactic structure of the sentence stem upon encountering the unpredictable target word.

In Lesage et al. (2014), such confounds are avoided by sufficient jittering of the delay preceding the first context sentence and the second sentence stem. However, the content of the second sentence stem (the event of interest) reflects the predictions generated on the grounds of the first sentence context but also predicts the occurrence of the target word of the second sentence. The next section will make the case that these findings may reflect the modulation of hemodynamic activity by task difficulty, which is only confounded with predictability.

6.2.2. Networks: task-positive or task-negative?

Moberget et al*.* (2014) employed scrambled sentence stems as control contexts with minimal predictability for the final target word. In fact, the authors argued that, in studies similarly contrasting scrambled sentences with semantically and syntactically canonical sentences, 'the critical variable was the predictability of the presented words– with increased cerebellar involvement for predictable relative to unpredictable conditions.' (*ibid*, p. 2876). Yet scrambled sentences do not provide an unproblematic baseline if prediction per se is addressed in contrasting 'Congruent > Scrambled'. One major source of sentence comprehension difficulty is the interference of material that partially matches syntactic and semantic retrieval cues (see Glaser, Martin, Van Dyke, Hamilton, & Tan (2013) for discussion). Moreover, such anomalous sentences may engage the semantic integration process to a greater degree in participants' attempt to comprehend the sentence (Brown & Hagoort, 1993). For example, syntactic violations traditionally activate areas involved in syntactic processing, since structure building, agreement checking and other putative operations are disrupted, and hence extra attention is paid to these aspects (Kaan & Swaab, 2002). Greater activations often occur, for instance, in the left inferior frontal gyrus for violated sentences than for normal sentences (see discussion in Hagoort et al., 2009), and indeed scrambled sentences involve syntactic and semantic violations to a larger extent than sentences with no such violations. Cognitive control and conflict monitoring are arguably engaged more in scrambled sentences as compared to the other sentence types. Maintenance-related processes are also involved in keeping information accessible whenever the currently described events seize to be consistent with the global situation model under construction at each stage. With respect to sentence comprehension, then, scrambled sentences are arguably more taxing than congruent ones. This would mean that the ‘Congruent > Scrambled’ contrast would be reflecting not (just) ‘Predictable > Not predictable’, but also a more generic ‘Low Demands > High Demands’ discrepancy, owed to the larger demands in semantic and syntactic integration, cognitive control, conflict monitoring and more attempts for reanalysis made in the ‘Scrambled’ as compared to the ‘Congruent’ condition.

How then would a ‘Low Demands > High Demands’ contrast yield this activation in the posterolateral regions of HVIIa Crus I/II? Functional connectivity studies make clear that lobule (H)VIIa does not form a functionally unitary locus. While findings in different studies are not unequivocal, they show that there are distinct regions within this lobule that are functionally connected with different prefrontal cortical areas. In particular, they demonstrate that, apart from the well-established involvement of regions in (H)VIIa in the executive control network, there are distinct regions within this lobule that form crucial hubs of the default-mode network. These regions occupy a sizeable posterolateral portion of HVIIa Crus I/II (Krienen & Buckner, 2009; Buckner et al., 2011; Bernard et al., 2012; for discussion on the absence of such findings in Habas et al. (2009), see Buckner et al. (2011), p. 2324 and p. 2340). Hubs of the default-mode network consistently demonstrate increased activity during rest or low-demand tasks as opposed to high-demand tasks across a broad range of paradigms, and have been argued to support internal self-reflective thought. The suppression of this network in attention-demanding tasks is manifested in the form of ‘task-induced deactivations’, traditionally yielding significant activation in contrasts of ‘ Low Demands > High Demands’, given their larger negative BOLD signal amplitude in conditions posing higher demands (e.g. Mazoyer, Zago, Mellet, Bricogne, Etard, Houde, Crivello, Joliot, Petit, & Tzourio-Mazoyer, 2001; Binder, Frost, Hammeke, Bellgowan, Rao, & Cox, 1999; Shulman, Fiez, Corbetta, Buckner, Miezin, Raichle, & Petersen, 1997; McKiernan, Kaufman, Kucera-Thompson, & Binder, 2003; Gusnard & Raichle, 2001; Raichle, MacLeod, Snyder, Powers, Gusnard, & Shulman, 2001; Raichle & Gusnard, 2005).

There is indeed substantial evidence for the task-induced deactivation of these specific posterolateral regions of HVIIa Crus I/II, a fortiori in non-linguistic tasks and with no apparent demand for predictive processing. In one such study, participants were required to indicate whether a visually presented digit-symbol probe-pair was present or absent in an array of nine digit-symbol probe-pairs. Along with the deactivation of other major hubs of the default-mode network, the data also demonstrated strong deactivations in posterior regions of right lobule HVIIa Crus I/II (Rao, Motes, & Rypma, 2014). Elsewhere, Harrison et al. (2008) studied correlated activity fluctuations of the default mode network regions during three conditions: rest with eyes closed; a moral dilemma task; a Stroop task. They showed a striking uniformity in the anatomy of the default mode network across these conditions. Crucially, one of its hubs was a posterolateral region of HVIIa Crus I, showing activation in the resting state and moral dilemma, but deactivation in the Stroop task. Fransson (2005) also reports that one of the brain regions that correlated positively with other hubs of this network (precuneus / posterior cingulate cortex) during resting state and in a condition with eyes closed was the left and right posterolateral region of HVIIa Crus I. In another study (Fransson, 2006), one such node that showed significantly more activity during rest compared to a two-back working memory task was represented by a substantial cluster in posterolateral regions of the right HVIIa Crus I . Gusnard, Akbudak, Shulman, and Raichle (2001) examined hemodynamic activity related with judgments in an internally cued and an externally cued condition. Along with medial prefrontal cortical activations, the contrast of ‘Internally > Externally cued judgments’ yielded an activation in the posterolateral region of right HVIIa Crus I/II. Moreover, the HVIIa Crus I/II peak activation coordinates in these contrasts are markedly adjacent to those found in Krienen and Buckner (2009) as peak coordinates of functional connectivity with their medial prefrontal cortical map- a major hub of the default mode network. The involvement of these regions in the default-mode network need not negate their computational properties in implementing internal models. It may simply suggest that internal models in these regions support processes of the default-mode network. Table 2 illustrates the adjacency of the peak activation coordinates from Moberget et al. (2014; ‘Congruent > Scrambled’) with those reported in studies on the engagement and disengagement of the default-mode network. Highly consistent with this interpretation is also the activation that Moberget et al. (2014) report in lobule IX for both ‘Congruent > Scrambled’ and ‘Incongruent > Scrambled’, as lobule IX is also part of the default mode network (Habas et al., 2009).

[Please insert table 2 here]

Likewise, in Lesage, Hansen, and Miall (2014), the parametric modulation by cloze probability in a right posterolateral region of HVIIa Crus II may not exclusively pertain to predictability. Lower cloze probability entails higher attentional and verbal working memory demands, as the memory trace of preceding linguistic information needs to be refreshed in order to check against a range of different completions. Indeed, the magnitude of deactivation of the default-mode network increases in a graded fashion in accordance with task load and error-proneness (McKiernan et al., 2003; Singh & Fawcett, 2008; Polli, Barton, Cain, Thakkar, Rauch, & Manoach, 2005; Persson, Lustig, Nelson, & Reuter-Lorenz, 2007; Kincses, Johansen-Berg, Tomassini, Bosnell, Matthews, & Beckmann, 2008; Esposito, Bertolino, Scarabino, Latorre, Blasi, Popolizio, Tedeschi, Cirillo, Goebel, & Di Salle, 2006). Hence the modulation of activity in this posterolateral region of right HVIIa Crus II by predictability may reflect the modulation of the disengagement of the default mode network by task difficulty. This would be a by-product of predictive processing, which may instead be carried out by cerebral cortical regions.

6.3. *Directions for improvements*

The discussion here has made clear the need for a number of improvements in the methods and considerations on the different findings. For both TMS and fMRI studies, a substantial improvement would address the need for carefully designed stimulus sets, whereby distinct experimental conditions would be exclusively matched in all other aspects apart from those of interest, i.e. the predictability of upcoming stimuli. In any other case, modulation of cerebellar engagement by predictability may be confounded with other mechanisms which cerebellar circuitry may also support, such as those in processing action-related semantics (e.g. Lesage et al., 2012). Both TMS and fMRI studies should also take into account the functional connectivity of different regions within (H)VIIa Crus I/II with distinct prefrontal cortical areas. While a great part of this lobule is embedded in the executive control network, the existence of default-mode network hubs in certain posterolateral regions of HVIIa Crus I/II (e.g. Buckner et al., 2011; Bernard et al., 2012) suggests that these regions may show modulation by task difficulty as other hubs of this network in the cerebral cortex (e.g. Polli et al., 2007; Singh & Fawcett, 2008). This should be taken into account when interpreting the modulation of activity in lobule (H)VIIa by predictability in language processing, especially when activations occur within these posterolateral regions of lobule HVIIa (e.g. Moberget et al., 2014). Future studies would thus benefit from the use of further control conditions and/or tasks that concomitantly vary with respect to processing demands. These would help establish whether modulation of cerebellar activity by predictability is reduced to generic modulation by task load. Alternatively, future studies should show that linguistic predictability modulates activity in regions other than those that non-linguistic predictability does, or that different types of linguistic prediction (e.g. phrasal, semantic) yield activations in distinct cerebellar lobular regions. For TMS studies in particular, double-cone coils should be used to provide appreciable strength in the stimulation of the deeper cerebellar hemispheres (Lesage et al., 2012; Hardwick, Lesage, & Miall, 2014), as compared to flat figure-of-eight-shaped coils (Argyropoulos & Muggleton, 2013). On the other hand, the focality of the double-cone coil is limited as compared to that of a small figure-of-eight-shaped coil. Nevertheless, the stimulation of adjacent, motor-projecting lobules need not be an insurmountable obstacle, provided that the different experimental conditions are well-matched with respect to properties to which motor-projecting lobules may be sensitive. While lobule VIIa also includes a vermal component, known as the ‘neocerebellar vermis’, the limited size of this portion and its adjacency to paravermal regions makes it difficult to stimulate in a selective fashion. This becomes an issue, when considering the involvement of paravermal regions in articulatory processes (Petersen et al., 1989; Frings et al., 2006). Irrespective of the coil shape and size, neuronavigated localization of the stimulation site (e.g. Halko et al., 2014) would contribute significantly towards identifying the particular lobular regions of (H)VIIa that receive the greatest stimulation.

7. *Outstanding questions*

Aside from overcoming the aforementioned weaknesses in methods and the interpretations of the findings yielded so far, this newly emerging field of studies on cerebellar contributions to ‘non-motor’ aspects of predictive language processing would benefit substantially from addressing a series of outstanding questions. Some of those are outlined below:

1. *What is the connectivity?* In control-theoretic accounts of cerebellar internal models, the anatomical evidence for the connectivity of specific cerebellar lobules with specific cerebral cortical areas is available for both motor control

as well as higher cognition (Kelly & Strick, 2003). Discussions of cerebellar contributions to non-motor aspects of language processing have piggy-backed on the latter. Many authors have speculated that input to the cerebellum from Broca's area would provide an efference copy based on which the cerebellum would generate and transmit its predictions back to the prefrontal cortex (Ito, 2000a; Argyropoulos, 2009; Lesage et al., 2012). Functional connectivity studies (e.g. Buckner et al., 2011) are consistent with those ideas. However, apart from evidence on the cortico-pontine projection from area 45B (Schmahmann & Pandya, 1997) and the fact that such projections have massively expanded in humans (Ramnani et al., 2006), anatomical evidence for cerebello-dentate projections back to area 45B or for temporo-parietal cortical projections to the inferior olive remains poor (see Ito (2008) for some references).

2. *What about mirror neurons?* What is the relationship between cerebellar internal models and alleged cerebral cortical internal models that mirror neuron circuits implement (e.g. Iacoboni, 2005; Hurley, 2008)? Do cerebellar internal models output predictions by covertly employing the production system in perceiving conspecifics (e.g. Blakemore & Decety, 2001), like the alleged internal models in mirror-neurons? If so, what are the consequences for interpreting cerebellar involvement in language comprehension? Little attention has been paid to these questions so far (Miall, 2003; Strick, Dum, & Fiez, 2009).
3. *Is there a level-specific topography in the linguistic predictions of the cerebellum?* If the cerebellum is involved in predictive language processing after all, do different lobules or lobular regions generate predictions for different levels (phonology, semantics)? No study has so far identified the particular level at which these predictions occur. The effect of stimulating the posterior vermis/paravermis on phrasal predictions (Argyropoulos, 2011a) could in principle be contrasted to the effects of lateral cerebellar stimulation on semantic predictions (Lesage, 2012; Argyropoulos & Muggleton, 2013), but there are major inherent weaknesses in those studies to establish this.
4. *What about production?* The studies conducted so far have all addressed the involvement of cerebellar forward models in 'non-motor' aspects of prediction in language comprehension. This has been mainly driven by considerations of

confounds with articulatory motor control that studying language generation would involve. However, prediction may involve both forward and inverse internal models at different levels of both language comprehension and generation (Pickering & Garrod, 2013). Indeed, the cerebellum may encode both forward and inverse internal models (Wolpert & Kawato, 1998). Can we then study cerebellar predictions in 'non-motor' aspects of language production?

8. *Conclusion*

Recent findings on the contributions of neocerebellar circuitry to the 'non-motor' aspects of predictive language processing are highly consistent with the involvement of the cerebellum in 'higher cognitive' aspects of behavior as well as with the multi-modal nature of this unitary cerebellar computation. However, a series of significant limitations in the few studies conducted so far raise questions about the interpretation of their findings. Addressing those weaknesses will provide the opportunity to investigate whether cerebellar internal models play a role in 'non-motor' aspects of predictive language processing. Above all, such research will improve our understanding of the 'modulatory' role of the cerebellum in language processing from a computationally grounded perspective that integrates motor, cognitive and affective aspects of cerebellar function (Mariën et al., 2014).

References

Ackermann, H., Wildgruber, D., & Grodd, W. (1998). Does the cerebellum contribute to cognitive aspects of speech production? A functional MRI study in humans. *Neuroscience Letters*, *247*, 187-190.

Ackermann, H., Mathiak, K., & Ivry, R. B. (2004). Temporal organization of —internal speech as a basis for cerebellar modulation of cognitive functions. *Behavioral and Cognitive Neuroscience Reviews*, *3*, 14-22.

Ackermann, H., Mathiak, K., & Riecker, A. (2007). The contribution of the cerebellum to speech production and speech perception: Clinical and functional imaging data. *Cerebellum*, *6*, 202–213.

Albus, J.S. (1971). A theory of cerebellar function. *Mathematical Biosciences*, *10*, 25– 61.

Altmann, G.T.M., & Kamide, Y. (1999). Incremental interpretation at verbs. *Cognition*, *73*, 247–264.

Arasanz, C.P., Staines, W.R., Roy, E.A., & Schweizer, T.A. (2012). The cerebellum and its role in word generation: a cTBS study. *Cortex*, *48*(6), 718-724.

Argyropoulos, G. P. (2008). The subcortical foundations of grammaticalization. In A. D. M. Smith, K. Smith, & R. Ferrer i Cancho (Eds.). *The evolution of language: Proceedings of the 7th international conference on the evolution of language* (pp. 10–17). Singapore: World Scientific.

Argyropoulos, G. P. (2009). Neocerebellar emulation in language processing. In K. Alter, M. Horne, M. Lindgren, M. Roll, & J. von Koss Torkildsen (Eds.). *Brain Talk: Discourse with and in the brain. Papers from the first Birgit Rausing language program conference in linguistics* (pp. 193–206). Lund: Lund University, Media Tryck.

Argyropoulos, G. P. (2011a). Cerebellar theta-burst stimulation selectively enhances lexical associative priming. *Cerebellum*, *10*(3), 540-50.

Argyropoulos, G.P. (2011b). *The neocerebellar Kalman filter linguistic processor: from grammaticalization to transcranial magnetic stimulation.* Unpublished doctoral dissertation. University of Edinburgh.

https://www.era.lib.ed.ac.uk/bitstream/1842/5694/2/Argyropoulos%202011.pdf

Argyropoulos, G. P., Kimiskidis, V., & Papagiannopoulos, S. (2011). Theta-burst stimulation of the right neocerebellar vermis selectively disrupts the practice-

induced acceleration of lexical decisions. *Behavioral Neuroscience*, *125*(5), 724-34.

Argyropoulos, G.P., & Muggleton, N. (2013). Effects of cerebellar stimulation on processing semantic associations. *Cerebellum*, *12*(1), 83-96.

Balsters, J.H., Cussans, E., Diedrichsen, J., Phillips, K.A., Preuss, T.M., Rilling, J.K., & Ramnani, N. (2010). Evolution of the cerebellar cortex: the selective expansion of prefrontal-projecting cerebellar lobules. *Neuroimage*, *49*, 2045-2052

Balsters, J. H., & Ramnani, N. (2011). Cerebellar Plasticity and Automation of Rule-Related Processing *Journal of Neuroscience*, *31*(6), 2305–2312.

Balsters, J. H., & Ramnani, N. (2008). Symbolic representations of action in the human cerebellum. *Neuroimage 43*(2), 388-398.

Balsters, J.H., Whelan, C.D., Robertson, I.H., & Ramnani, N. (2013). Cerebellum and cognition: evidence for the encoding of higher order rules. *Cerebral Cortex*, *23*(6), 433-43.

Bellebaum, C., & Daum, I. (2011). Mechanisms of cerebellar involvement in associative learning. *Cortex*, *47*(1), 128–36.

Bernard, J.A., Seidler, R.D., Hassevoort, K.M., Benson, B.L., Welsh, R.C., Wiggins, J.L., Jaeggi, S.M., Buschkuehl, M., Monk, C.S., Jonides, J., & Peltier, S.J. (2012). Resting state cortico-cerebellar functional connectivity networks: a comparison of anatomical and self-organizing map approaches. *Frontiers in Neuroanatomy*, *6*(31).

Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S., Rao, S. M., & Cox, R. W. (1999). Conceptual processing during the conscious resting state. A functional MRI study. *Journal of Cognitive Neuroscience*,*11*, 80–95.

Blakemore, S.-J., & Decety, J. (2001). From the perception of action to the understanding of intention. *Neuroscience*, *2*, 561-567.

Blakemore, S.-J., Frith, C. D., & Wolpert, D. M. (2001). The cerebellum is involved in predicting the sensory consequences of action. *NeuroReport*, *12*, 1879-84.

Bloedel, J.R. (1992). Functional heterogeneity with structural homogeneity - How does the cerebellum operate? *Behavioral and Brain Sciences*, *15*, 666–678.

Blumstein, S. E., Milberg, W. P., & Shrier, R. (1982). Semantic processing in aphasia: Evidence from an auditory lexical decision task. *Brain and Language*, *17*, 301–315.

Brown, C., & Hagoort, P. (1993). The processing nature of the N400: Evidence from

masked priming. *Journal of Cognitive Neuroscience*, *5*, 34 –44.

Brysbaert, M., Warriner, A.B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, *46*, 904-911.

Buccino, G., Riggio, L., Melli, G., Binkofski, F., Gallese, V., & Rizzolatti, G. (2005). Listening to action-related sentences modulates the activity of the motor system: a combined TMS and behavioral study. *Brain Research Cognitive Brain Research*, *24*, 355–63.

Buckner, R.L., Krienen, F.M., Castellanos, A., Diaz, J.C., & Yeo, B.T.T. (2011). The organization of the human cerebellum estimated by intrinsic functional connectivity *Journal of Neurophysiology, 106*, 2322–2345.

Callan, D.E., Kawato, M., Parsons, L., & Turner, R. (2007). Speech and song: the role of the cerebellum. *Cerebellum*, *6*, 321-327.

Canavan, A. G., Sprengelmeyer, R., Diener, H. C., & Hömberg, V. (1994). Conditional associative learning is impaired in cerebellar disease in humans. *Behavioral Neuroscience*, *108*, 475-85.

Cappa, S.F., Sandrini, M., Rossini, P.M., Sosta, K., & Miniussi, C. (2002). The role of the left frontal lobe in action naming: rTMS evidence. *Neurology*, *59*, 720–3.

Chomsky, N. (1981). *Lectures on government and binding.* Dordrecht: Foris.

Christian, K. M., & Thompson, R. F. (2005). Long-term storage of an associative memory trace in the cerebellum. *Behavioral Neuroscience*, *119*, 526-37.

Coffman, K., Dum, R. P., & Strick, P. L. (2011). Cerebellar vermis is a target of projections from the motor areas in the cerebral cortex. *Proceedings of the National Academy of Sciences*, *108*(38), 16068–73.

Coltheart, M. (1981). The MRC Psycholinguistic Database. *Quarterly Journal of Experimental Psychology*, *33*A, 497-505.

Courchesne, E., & Allen, G. (1997). Prediction and preparation, fundamental functions of the cerebellum. *Learning & Memory*, *4*(1), 1-35.

Daskalakis, Z.J., Paradiso, G.O., Christensen, B.K., Fitzgerald, P.B., Gunraj, C., & Chen, R. (2004). Exploring the connectivity between the cerebellum and motor cortex in humans. *Journal of Physiology*, *557*, 689–700.

Daum, I., Ackermann, H., Schugens, M.M., Reimold, C., Dichgans, J., & Birbaumer, N. (1993). The cerebellum and cognitive functions in humans. *Behavioral Neuroscience*, *107*(3), 411–19.

De Smet, H.J., Paquier, P., Verhoeven, J., & Mariën, P. (2013). The cerebellum: its role in language and related cognitive and affective functions. *Brain & Language*, *127*, 334-342.

Desmond, J. E., & Fiez, J. A. (1998). Neuroimaging studies of the cerebellum: Language, learning, and memory. *Trends in Cognitive Sciences*, *2*(9), 355-62.

Desmond, J. E., Gabrieli, J. D. E., & Glover, G. H. (1998). Dissociation of frontal and cerebellar activity in a cognitive task: evidence for a distinction between selection and research. *Neuroimage*, *7*, 368-76.

Desmurget, M., Gréa, H., Grethe, J. S., Prablanc, C., Alexander, G. E., & Grafton, S. T. (2001). Functional anatomy of nonvisual feedback loops during reaching: a positron emission tomography study. *Journal of Neuroscience*, *21*, 2919-2928.

Devlin, J. T., & Watkins, K. E. (2007). Stimulating language: insights from TMS. *Brain*, *130*(3), 610-622.

Ding, H., Qin, W., Jiang, T., Zhang, Y., & Yu, C. (2012). Volumetric variation in subregions of the cerebellum correlates with working memory performance. *Neuroscience Letters*, *508*(1), 47–51.

Dow, R. S., & Moruzzi, G. (1958). *The physiology and pathology of the cerebellum*. Minneapolis: University of Minnesota Press.

Drepper, J., Timmann, D., Kolb, F. P., & Diener, H. C. (1999). Non-motor associative learning in patients with isolated degenerative cerebellar disease. *Brain*, *122*, 87-97.

E, K.-H., Chen, S.-H. A., Ho, M.-H. R., & Desmond, J. E. (2014). A meta-analysis of cerebellar contributions to higher cognition from PET and fMRI studies. *Human Brain Mapping*, *35*(2), 593-615.

Esposito, F., Bertolino, A., Scarabino, T., Latorre, V., Blasi, G., Popolizio, T., Tedeschi, G., Cirillo, S., Goebel, R., & Di Salle, F. (2006). Independent component model of the default-mode brain function: Assessing the impact of active thinking. *Brain Research Bulletin, 70*, 263–269.

Fedorenko, E., Hsieh, P.-J., Nieto-Castañón, A., Whitfield-Gabrieli, S., & Kanwisher, N. (2010). New method for fMRI investigations of language: defining ROIs functionally in individual subjects. *Journal of Neurophysiology*, *104*(2), 1177–94.

Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology*, *47*, 164-203.

Ferrucci, R., Marceglia, S., Vergari, M., Cogiamanian, F., Mrakic-Sposta, S., Mameli,

F., Zago, S., Barbieri, S., & Priori, A. (2008). Cerebellar transcranial direct current stimulation impairs the practice-dependent proficiency increase in working memory. *Journal of Cognitive Neuroscience*, *20*(9), 1687-97.

Fiez, J. A., & Raichle, M. (1997). Linguistic processing. *International Review of Neurobiology, 41*, 233-54.

Fiez, J.A., Petersen, S.E., Cheney, M.K., & Raichle, M.E. (1992). Impaired nonmotor learning and error detection associated with cerebellar damage. A single case study. *Brain*, *115*, 155–78.

Fransson, P. (2005). Spontaneous low-frequency BOLD signal fluctuations: an fMRI investigation of the resting-state default mode of brain function hypothesis. *Human Brain Mapping, 26*, 15–29.

Fransson, P. (2006). How default is the default mode of brain function? further evidence from intrinsic BOLD signal fluctuations. *Neuropsychologia*, *44*, 2836–2845.

Friederici, A. D. (2012). The cortical language circuit: from auditory perception to sentence comprehension. *Trends in Cognitive Sciences*, *16*(5), 262–8.

Frings, M., Dimitrova, A., Schorn, C.F., Elles, H.-G., Hein-Kropp, C., Gizewski, E.R., Diener, H.C., & Timmann, D. (2006). Cerebellar involvement in verb generation: an fMRI study. *Neuroscience Letters*, *409*, 19–23.

Gebhart, A. L., Petersen, S. E., & Thach, W. T. (2002). Role of the posterolateral cerebellum in language. *Annals of the New York Academy of Science*, *978*, 318-33.

Ghosh S.S., Tourville, J.A., & Guenther F.H. (2009). A neuroimaging study of premotor lateralization and cerebellar involvement in the production of phonemes and syllables. *Journal of Speech, Language, and Hearing Research*, *51*(5), 1183-1202.

Gilbert, P.F. & Thach, W.T. (1977). Purkinje cell activity during motor learning. *Brain Research*, *128*, 309–328.

Glaser, Y.G., Martin, R.C., Van Dyke, J.A., Hamilton, A.C., & Tan, Y. (2013). Neural basis of semantic and syntactic interference in sentence comprehension. *Brain & Language*, *126*, 314–326.

Golfinopoulos, E., Tourville, J. a, Bohland, J. W., Ghosh, S. S., Nieto-Castanon, A., & Guenther, F. H. (2011). fMRI investigation of unexpected somatosensory feedback perturbation during speech. *Neuroimage*, *55*(3), 1324–38.

Golfinopoulos, E., Tourville, J., & Guenther, F. H. (2010). The integration of large-scale neural network modeling and functional brain imaging in speech motor control. *Neuroimage*, *52*(3), 862–74.

Greicius, M.D., & Menon, V. (2004). Default-mode activity during a passive sensor task: Uncoupled from deactivation but impacting activation. *Journal of Cognitive Neuroscience*, *16*, 1484–1492.

Grimaldi, G., Argyropoulos, G.P., Bastian, A., Cortes, M., Davis, N.J., Edwards, D., Ferrucci, R., Fregni, F., Galea, J.M., Hamada, M., Manto, M., Miall, R.C., Morales-Quezada, L., Pope, P.A., Priori, A., Rothwell, J., Tomlinson, S.P., & Celnik, P. (2014). Cerebellar transcranial direct current stimulation (ctDCS): a novel approach to understand cerebellar function in health and disease. *Neuroscientist* [epub ahead of print].

Grimaldi, G., Argyropoulos, G.P., Boehringer, A., Celnik, P., Edwards, M.J., Ferrucci, R., Galea, J.M., Groiss, S.J., Hiraoka, K., Kassavetis, P., Lesage, E., Manto, M., Miall, R.C., Priori, A., Sadnicka, A., Ugawa, Y., & Ziemann, U. (2013). Non-invasive cerebellar stimulation—a consensus paper. *Cerebellum*, *13*(1), 121-138.

Grush, R. (2004). The emulation theory of representation: motor control, imagery, and perception. *Behavioral and Brain Sciences*, *27*, 377-435.

Guediche, S., Holt, L. L., Laurent, P., Lim, S.-J., & Fiez, J. A. (2014). Evidence for cerebellar contributions to adaptive plasticity in speech perception. *Cerebral Cortex*. [Epub ahead of print].

Gusnard, D.A., & Raichle, M.E. (2001). Searching for a baseline: functional imaging and the resting human brain. *Nature Reviews Neuroscience, 2*, 685–694.

Gusnard, D.A., Akbudak, E., Shulman, G.L., & Raichle, M.E. (2001). Medial prefrontal cortex and self-referential mental activity: relation to a default mode of brain function. *Proceedings of the National Academic of Sciences*, 98(7), 4259–4264.

Habas, C., Kamdar, N., Nguyen, D., Prater, K., Beckmann, C. F., Menon, V., & Greicius, M. D. (2009). Distinct cerebellar contributions to intrinsic connectivity networks. *Journal of Neuroscience*, *29*(26), 8586–94.

Hagoort, P., Baggio, G., & Willems, R. M. (2009). Semantic unification. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences*, 4th ed. (pp. 819-836). Cambridge, MA: MIT Press.

Halko, M.A., Farzan, F., Eldaief, M.C., Schmahmann, J.D., & Pascual-Leone, A.

(2014). Intermittent theta-burst stimulation of the lateral cerebellum increases functional connectivity of the default network. *Journal of Neuroscience, 34*(36), 12049-12056.

Hanakawa, T., Dimyan, M.A., & Hallett, M. (2008). Motor planning, imagery, and execution in a distributed motor network: a time-course study with functional MRI. *Cerebral Cortex*, *18*, 2775–88.

Hardwick, R.M., Lesage, E., & Miall, R.C. (2014). Cerebellar transcranial magnetic stimulation: the role of coil geometry and tissue depth. *Brain Stimulation*, *7*, 643-649.

Harrison, B.J., Pujol, J., López-Solà, M., Hernández-Ribas, R., Deus, J., Ortiz, H., Soriano-Mas, C., Yücel, M., Pantelis, C., & Cardoner, N. (2008). Consistency and functional specialization in the default mode brain network. *Proceedings of the National Academic of Sciences*, *105*(28), 9781–9786.

Hashimoto, M., & Ohtsuka, K. (1995). Transcranial magnetic stimulation over the posterior cerebellum during visually guided saccades in man. *Brain*, *118*, 1185–1193.

Hauk, O., Johnsrude, I., & Pulvermuller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron*, *41*, 301–7.

Heinks-Maldonado, T. H., Nagarajan, S. S., & Houde, J. F. (2006). Magnetoencephalographic evidence for a precise forward model in speech production. *Neuroreport*, *17*(13), 1375–1379.

Hickok, G. (2012). Computational neuroanatomy of speech production. *Nature Reviews Neuroscience*, *13*, 135–145.

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, *8*, 393–402.

Hosemann, J., Herrmann, A., Steinbach, M., Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2013). Lexical prediction via forward models: N400 evidence from German Sign Language *Neuropsychologia*, *51*, 2224–2237.

Hurley, S. (2008). The shared circuits model (SCM): How control, mirroring, and simulation can enable imitation, deliberation, and mindreading. *Behavioral and Brain Sciences*, 31, 1- 58.

Hutchison, K. A. (2003). Is semantic priming due to association strength or feature overlap? A *micro*analytic review. *Psychonomic Bulletin & Review*, *10*(4), 785-813.

Iacoboni, M. (2008). The role of premotor cortex in speech perception: evidence from

fMRI and rTMS. *Journal of Physiology*, *102*(1-3), 31-34.

Imamizu, H, Miyauchi, S., Tamada, T., Sasaki, Y., Takino, R., Pütz, B., Yoshioka, T., & Kawato, M. (2000). Human cerebellar activity reflecting an acquired internal model of a new tool. *Nature*, *403*, 192-5.

Ito, M. (1970). Neurophysiological basis of the cerebellar motor control system. *International Journal of Neurology*, *7*, 162-76.

Ito, M. (1984). *The cerebellum and neural control.* New York: Raven Press.

Ito, M. (2000a). Neural control of cognition and language. In A. Marantz, Y. Miyashita, & W. O'Neil (Eds.), *Image, language, brain* (pp. 149-62). Cambridge, Massachusetts: MIT Press.

Ito, M. (2000b). Neurobiology: internal model visualized. *Nature*, *403*, 153-154.

Ito, M. (2008). Control of mental activities by internal models in the cerebellum. *Nature Reviews Neuroscience*, *9*, 304-13.

Jones, L.L., & Estes, Z. (2012). Lexical Priming: associative, semantic, and thematic influences on word recognition. In: J. Adelman (Ed.), *Visual Word Recognition, Vol II: Meaning and Context, Individuals and Development* (pp. 44-72). Psychology Press.

Jordan, M. I., & Wolpert, D. M. (2000). Computational motor control. In M. S. Gazzaniga, (Ed.), *The new cognitive neurosciences* (pp. 601-618). Cambridge, MA: MIT Press.

Jueptner, M., Frith, C.D., Brooks, D.J., Frackowiak, R.S.J., & Passingham, R.E. (1997). Anatomy of motor learning. II. Subcortical structures and learning by trial and error. *Journal of Neurophysiology*, *77*, 1325-37.

Kaan, E., & Swaab, T.Y. (2002). The brain circuitry of syntactic comprehension. *Trends in Neurosciences*, *6*(8), 350-356.

Kawato, M. (1999). Internal models for motor control and trajectory planning, *Current Opinion in Neurobiology*, *9*, 718-27.

Kawato, M., Kuroda, T., Imamizu, H., Nakano, E., Miyauchi, S., & Yoshioka, T. (2003). Internal forward models in the cerebellum: fMRI study on grip force and load force coupling. *Progress in Brain Research*, *142*, 171-88.

Kelly, R.M., & Strick, P.L. (2003). Cerebellar loops with motor cortex and prefrontal cortex of a nonhuman primate. *Journal of Neuroscience*, *23(23)*, 8432–44.

Kincses, Z.T., Johansen-Berg, H., Tomassini, V., Bosnell, R., Matthews, P.M., & Beckmann, C.F. (2008). Model-free characterization of brain functional networks

for motor sequence learning using fMRI. *Neuroimage*, *39*(4), 1950–1958.

Kotz, S.A., & Schwartze, M. (2010). Cortical speech processing unplugged: a timely subcortico-cortical framework. *Trends in Cognitive Sciences*, *14*, 392-399.

Krienen, F.M., & Buckner, R.L. (2009). Segregated fronto-cerebellar circuits revealed by intrinsic functional connectivity. *Cerebral cortex*, *19*, 2485-2497.

Lang, C. E., & Bastian, A. J. (2002). Cerebellar damage impairs automaticity of a recently practiced movement. *Journal of Neurophysiology*, *87*(3), 1336-47.

Leggio, M.G., Silveri, M.C., Petrosini, L., & Molinari, M. (2000). Phonological grouping is specifically affected in cerebellar patients: a verbal fluency study. *Journal of Neurology, Neurosurgery, and Psychiatry*, *69*, 102–6.

Leggio, M.G., Tedesco, A. M., Chiricozzi, F. R., Clausi, S., Orsini, A., & Molinari, M. (2008). Cognitive sequencing impairment in patients with focal or atrophic cerebellar damage. *Brain*, *131*, 1332-1343.

Leiner, H.C. (2010). Solving the mystery of the human cerebellum. *Neuropsychology Review*, *20*, 229-235.

Lesage, E. (2013). *On the cerebellum and language: neurostimulation and imaging studies.* Unpublished doctoral dissertation. University of Birmingham. http://etheses.bham.ac.uk/4926/1/Lesage14PhD.pdf

Lesage, E., Hansen, P.C., & Miall, C.C. (2014). Cerebellar BOLD response to linguistic stimuli is modulated by predictability. *Society for Neuroscience abstracts.*

Lesage, E., Morgan, B.E., Olson, A.C., Meyer, A.S., & Miall, R.C. (2012). Cerebellar rTMS disrupts predictive language processing, *Current Biology*, *22*(18), 794-795.

Lieberman, P. (1963). Some effects of semantic and grammatical context on the production and perception of speech. *Language and Speech*, *6*, 172-187.

Liuzzi, G., Freundlieb, N., Ridder, V., Hoppe, J., Heise, K, Zimerman, M., Dobel, C., Enriquez-Geppert, S., Gerloff, C., Zwitserlood, P., & Hummel, F.C. (2010). The involvement of the left motor cortex in learning of a novel action word lexicon. *Current Biology*, *20*(19), 1745-51.

Maddox, W.T., Aparicio, P., Marchant, N.L., & Ivry, R.B. (2005). Rule-based category learning is impaired in patients with Parkinson's disease but not in patients with cerebellar disorders. *Journal of Cognitive Neuroscience*, *17*(5), 707-23.

Manto, M., Bower, J.M., Conforto, A.B., Delgado-García, J.M., da Guarda, S.N., Gerwig, M., Habas, C., Hagura, N., Ivry, R.B., Mariën, P., Molinari, M., Naito, E., Nowak, D.A., Oulad Ben Taib, N., Pelisson, D., Tesche, C.D., Tilikete, C., &

Timmann, D. (2012). Consensus paper: roles of the cerebellum in motor control--the diversity of ideas on cerebellar involvement in movement. *Cerebellum*, *11*(2), 457-87.

Mariën, P., Ackermann, H., Adamaszek, M., Barwood, C.H.S., Beaton, A., Desmond, J., De Witte, E., Fawcett, A.J., Hertrich, I., Küper, M., Leggio, M., Marvel, C., Molinari, M., Murdoch, B.E., Nicolson, R.I., Schmahmann, J.D., Stoodley, C.J., Thürling, M., Timmann, D., Wouters, E., & Ziegler, W. (2014). Consensus paper: language and the cerebellum: an ongoing enigma. *Cerebellum*, *13*, 386–410.

Mariën, P., Engelborghs, S., Fabbro, F., & De Deyn, P. P. (2001). The lateralized linguistic cerebellum: a review and a new hypothesis. *Brain and Language*, *79*, 580-600.

Marr, D. (1969). A theory of cerebellar cortex. *Journal of Physiology*, *202*, 437– 470.

Mazoyer, B., Zago, L., Mellet, E., Bricogne, S., Etard, O., Houde, O., Crivello, F., Joliot, M., Petit, L., & Tzourio-Mazoyer, N. (2001). Cortical networks for working memory and executive functions sustain the conscious resting state in man. *Brain Research Bulletin*, *54*, 287–298.

McCormick, D. A, & Thompson, R. F. (1984). Cerebellum: essential involvement in the classically conditioned eyelid response. *Science*, *223*, 296-9.

McKiernan, K.A., Kaufman, J.N., Kucera-Thompson, J., & Binder, J.R. (2003). A parametric manipulation of factors affecting task-induced deactivation in functional neuroimaging. *Journal of Cognitive Neuroscience, 15*, 394–408.

McRae, K., Hare, M., Elman, J. L., & Ferretti, T. (2005). A basis for generating expectancies for verbs from nouns. *Memory and Cognition*, *33*(7), 1174-84.

Medina, J.F. & Lisberger, S.G. (2009). Encoding and decoding of learned smooth-pursuit eye movements in the floccular complex of the monkey cerebellum, *Journal of Neurophysiology*, *102*, 2039-2054.

Miall, R. C. (2003). Connecting mirror neurons and forward models. *NeuroReport*, *14*(16), 1-3.

Miall, R. C., & Christensen, L. O. D. (2004). The effect of rTMS over the cerebellum in normal human volunteers on peg-board movement performance. *Neuroscience Letters*, 371, 185-189.

Miall, R. C., Christensen, L. O. D., Cain, O., & Stanley, J. (2007). Disruption of state estimation in the human lateral cerebellum. *Public Library of Science Biology*, *5*(11), 2733-44.

Miall, R. C., & King, D. (2008). State estimation in the cerebellum. *Cerebellum*, *7*, 572-6.

Middleton, F.A. & Strick, P.L. (2000). Basal ganglia and cerebellar loops: motor and cognitive circuits. *Brain Research Reviews*, *31*, 236-250.

Middleton, F. A., & Strick, P. L. (2001). Cerebellar projections to the prefrontal cortex of the primate. *Journal of Neuroscience*, *21*(2), 700-12.

Milberg, W. P., & Blumstein, S. E. (1981). Lexical decision and aphasia: evidence for semantic processing. *Brain and Language*, *14*, 371–385.

Moberget, T., Gullesen, E.H., Andersson, S., Ivry, R.B., & Endestad, T. (2014). Generalized role for the cerebellum in encoding internal models: evidence from semantic processing. *Journal of Neuroscience*, *34*(8), 2871–2878.

Müller, F., & Dichgans, J. (1994). Dyscoordination of pinch and lift forces during grasp in patients with cerebellar lesions. *Experimental Brain Research*, *101*, 485-92.

Nestor, P., Valdman, O., Niznikiewicz, M., Spencer, K., McCarley, R., & Shenton, M. (2006). Word priming in schizophrenia: associational and semantic influences *Schizophrenia Research*, *82*(2), 139-42.

Nowak, D. A., Timmann, D., & Hermsdorfer, J. (2007). Dexterity in cerebellar agenesis. *Neuropsychology*, *45*, 696-703.

Oliveri, M., Koch, G., Torriero, S., & Caltagirone, C. (2005) Increased facilitation of the primary motor cortex following 1 Hz repetitive transcranial magnetic stimulation of the contralateral cerebellum in normal humans. *Neuroscience Letters*, *376*, 188–193.

Oliveri, M., Bonnì, S., Turriziani, P., Koch, G., Gerfo, E. L., Torriero, S., Vicario, C.M., Petrosini, L. & Caltagirone, C. (2009). Motor and linguistic linking of space and time in the cerebellum. *PloS One*, *4*(11), e7933.

Perkell, J. S., Matthies, M. L., Lane, H., Guenther, F. H., Wilhelms-Tricarico, R., Wozniak, J. & Guiod, P. (1997). Speech motor control: acoustic goals, saturation effects, auditory feedback and internal models. *Speech Communication*, *22*, 227-250.

Persson, J., Lustig, C., Nelson, J., & Reuter-Lorenz, P. (2007). Age differences in deactivation: a link to cognitive control? *Journal of Cognitive Neuroscience*, *19*, 1021–1032.

Petersen, S. E., Fox, P. T., Posner, M. L., Mintun, M., & Raichle, M. E. (1989).

Positron emission tomographic studies of the processing of single words. *Journal of Cognitive Neuroscience*, *1*, 153-70.

Picard, H., Amado, I., Mouchet-Mages, S., Olié, J.- P., & Krebs, M.-O. (2008). The role of the cerebellum in schizophrenia: an update of clinical, cognitive, and functional evidences. *Schizophrenia Bulletin*, *34*(1), 155-172.

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, *27*, 169-225.

Pickering, M. J., & Garrod, S. (2007). Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences*, *11*(3), 105-10.

Pickering, M. J., & Garrod, S. (2013). Forward models and their implications for production, comprehension, and dialogue. *Behavioral and Brain Sciences*, *36*(4), 377-392.

Poeppel, D., Emmorey, K., Hickok, G., & Pylkkänen, L. (2012). Towards a new neurobiology of language. *Journal of Neuroscience*, *32*,14125–14131.

Polli, F., Barton, J., Cain, M., Thakkar, K., Rauch, S., & Manoach, D. (2005). Rostral and dorsal anterior cingulate cortex make dissociable contributions during antisaccade error commission. *Proceedings of the National Academy of Science*, *102*, 15700–15705.

Pope, P., & Miall, R.C. (2012). Task-specific facilitation of cognition by cathodal transcranial direct current stimulation of the cerebellum. *Brain Stimulation*, *5*, 84-94.

Pulvermüller, F., Hauk, O., Nikulin, V.V., & Ilmoniemi, R.J. (2005). Functional links between motor and language systems. *European Journal of Neuroscience*, *21*, 793–7.

Puttemans, V., Wenderoth, N., & Swinnen, S.P. (2005). Changes in brain activation during the acquisition of a multifrequency bimanual coordination task: from the cognitive stage to advanced levels of automaticity. *Journal of Neuroscience, 25*, 4270–4278.

Raboyeau, G., Marie, N., Balduyck, S., Gros, H., Démonet, J. -F., & Cardebat, D. (2004). Lexical learning of the English language: a PET study in healthy French subjects. *Neuroimage*, *22*(4), 1808–18.

Raichle, M.E., & Gusnard, D.A. (2005). Intrinsic brain activity sets the stage for expression of motivated behaviour. *Journal of Comparative Neurology*, *493*(1), 167-176.

Raichle, M.E., MacLeod, A.M., Snyder, A.Z., Powers, W.J., Gusnard, D.A., & Shulman, G.L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences*, *98*, 676–682.

Ramnani, N. (2006). The primate cortico-cerebellar system: anatomy and function. *Nature Reviews Neuroscience*, *7*, 511–522.

Ramnani, N., Behrens, T.E., Johansen-Berg, H., Richter, M.C., Pinsk, M.A., Andersson, J.L., Rudebeck, P., Ciccarelli, O., Richter, W., Thompson, A.J., Gross, C.G., Robson, M.D., Kastner, S., & Matthews, P.M. The evolution of prefrontal inputs to the cortico-pontine system: diffusion imaging evidence from macaque monkeys and humans (2006). *Cerebral Cortex*, *16*(6), 811-8.

Rao, N.K., Motes, M.A., & Rypma, B. (2014). Investigating the neural bases for intra-subject cognitive efficiency changes using functional resonance imaging. *Frontiers in Human Neuroscience*, *8*, 840.

Riva, D., & Giorgi, C. (2000). The cerebellum contributes to higher functions: evidence from a series of children surgically treated for posterior fossa tumours. *Brain*, *123*, 1051-1061.

Rogers, S. L., & Friedman, R. B. (2008). The underlying mechanisms of semantic memory in Alzheimer’s disease and semantic dementia. *Neuropsychologia*, *46*, 12-21.

Roxbury, T., McMahon, K., & Copland, D.A. (2014). An fMRI study of concreteness effects in spoken word recognition. *Behavioral and Brain Functions*, *10*(34),1-14.

Schmahmann, J. D. (2000). The role of the cerebellum in affect and psychosis. *Journal of Neurolinguistics*, *13*, 189-214.

Schmahmann, J. D., & Pandya, D. N. (1997). The cerebrocerebellar system. *International Review of Neurobiology*, *41*, 31-60.

Shapiro, K.A., Pascual-Leone, A., Mottaghy, F.M., Gangitano, M., & Caramazza, A. (2001). Grammatical distinctions in the left frontal cortex. *Journal of Cognitive Neuroscience*, *13*, 713–20.

Shulman, G.L., Fiez, J.A., Corbetta, M., Buckner, R.L., Miezin, F.M., Raichle, M.E., & Petersen, S.E. (1997). Common blood flow changes across visual tasks: I. Increases in subcortical structures and cortex. *Journal of Cognitive Neuroscience*, *9*, 648–663.

Singh, K.D., & Fawcett, I.P. (2008). Transient and linearly graded deactivation of the human default-mode network by a visual detection task. *Neuroimage*, *41*,100-112.

Staub, A., & Clifton, C. Jr. (2006). Syntactic prediction in language comprehension: evidence from either…or. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(2), 425-436.

Stoodley, C.J., & Schmahmann, J.D. (2009). Functional topography in the human cerebellum: a meta-analysis of neuroimaging studies. *Neuroimage, 44*(2), 489-501.

Strick P.L., Dum, R.P., & Fiez, J.A. (2009). Cerebellum and nonmotor function. *Annual Review of Neuroscience*, *32*, 413–434.

Sturt, P., & Lombardo, V. (2005). Processing coordinated structures: incrementality and connectedness. *Cognitive Science*, *29*, 291-305.

Thach, W. T. (1997). Context-response linkage. *International Review of Neurobiology*, *41*, 599-611.

Thompson, R. F., Bao, S., Chen, L., Cipriano, B. D., Grethe, J. S., Kim, J. J., Thompson, J. K., Tracy, J. A., Weninger, M. S., & Krupa, D. J. (1997). Associative learning. *International Review of Neurobiology*, *41*, 152-89.

Thompson-Schill, S. L., Kurtz, K. J., & Gabrieli, J. D. E. (1998). Effects of semantic and associative relatedness on automatic priming. *Journal of Memory and Language*, *38*, 440-58.

Tian, X., & Poeppel, D. (2010). Mental imagery of speech and movement implicates the dynamics of internal forward models. *Frontiers in Psychology*, *1*, 1–23.

Timmann, D., Drepper, J., Calabrese, S., Bürgerhoff, K., Maschke, M., Kolb, F.P., Daum, I., & Diener, H. C. (2004). Use of sequence information in associative learning in control subjects and cerebellar patients. *Cerebellum*, 3, 75–82.

Timmann, D., Drepper, J., Frings, M., Maschke, M., Richter, S., Gerwig, M., & Kolb, F.P. (2010). The human cerebellum contributes to motor, emotional and cognitive associative learning. A review. *Cortex*. *46*, 845–57.

Timmann, D., Drepper, J., Maschke, M., Kolb, F.P., Böring, D., Thilmann, A.F., & Diener, H.C. (2002). Motor deficits cannot explain impaired cognitive associative learning in cerebellar patients. *Neuropsychologia*, *40*, 788–800.

Timmann, D., Kolb, F. P., Baier, C., Rijntjes, M., Mueller, S. P., Diener, H. C., & Weiller, C. (1996). Cerebellar activation during classical conditioning of the human flexion reflex: a PET study. *NeuroReport*, *7*, 2056-60.

Toni, I., Krams, M., Turner, R., & Passingham, R. E. (1998). The time course of changes during motor sequence learning: a whole-brain fMRI study. *Neuroimage*,

8, 50-61.

Tourville, J., Reilly, K. J., & Guenther, F. H. (2008). Neural mechanisms underlying auditory feedback control of speech. *Neuroimage*, *39*(3), 1429–43.

Townsend, D., & Bever, T. G. (2001). *Sentence comprehension: the integration of habits and rules.* Cambridge, Massachusetts: MIT Press.

Tucker, J., Harding, A. E., Jahanshahi, M., Nixon, P. D., Rushworth, M., Quinn, N. P., Thompson, P. D., & Passingham, R. E. (1996). Associative learning in patients with cerebellar ataxia. *Behavioral Neuroscience*, *110*, 1229-34.

Watkins, K., & Paus, T. (2004). Modulation of motor excitability during speech perception: the role of Broca’s area. *Journal of Cognitive Neuroscience*, *16*(6), 978-987.

Wilson, M., & Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psychological Bulletin*, *131*(3), 460-473

Wolpert, D.M., & Kawato, M. (1998). Multiple paired forward and inverse models for motor control. *Neural Networks*, *11*, 1317–1329.

Wolpert, D.M., Miall, R.C., & Kawato, M. (1998). Internal models in the cerebellum. *Trends in Cognitive Sciences*, *2*(9), 338–47.

Figure Legends

Figure 1. Neural implementation of internal (forward) models in the cerebellum; left: an internal model implemented in cerebellar lobules HV,HVI,HVIIb,HVIII interacting with the motor cortex; right: an internal model implemented in cerebellar lobule (H)VIIa Crus I/II interacting with the prefrontal cortex. Its organization is explicitly based on the well-studied cerebro-cerebellar circuitry and corresponding control-theoretic accounts of motor control (left). A copy of a process involving the interaction of the prefrontal cortex with another brain area is sent via the pontine nuclei to its interconnected cerebellar lobules. Predictions generated from those are transmitted from the Purkinje cells of the cerebellar cortex via the ventrolateral neodentate nucleus and the thalamus back to those prefrontal areas. Predicted and actual consequences of the process copied by these cerebellar lobules are compared in the inferior olive. Any discrepancies between the two are fed via climbing fibers to the cerebellar cortex in the form of an error signal. Long-term depression is triggered at the parallel fiber-to-Purkinje cell synapses, updating the internal model. RN: red nucleus. Modified with permission from Macmillan Publishers Ltd: NATURE REVIEWS NEUROSCIENCE (Ramnani, N. (2006). The primate cortico-cerebellar system: anatomy and function. Nat. Rev. Neurosci., 7, 511–522), copyright (2006).

Figure 2. Schematic representation of an internal (forward) model involved in the comprehension of the sentence ‘Harry went out to fly his red flag’ for three different levels of prediction (Pickering & Garrod, 2007). This ‘emulator’ (Grush, 2004) is controlled by feedback from a Kalman filter, weighing predictions against analysis of

the speech input at each (arbitrarily defined here) step. In case of strong predictions and/or noisy inputs, there is ‘low Kalman gain’ (strong top-down influence on the analysis); in case of poor predictions and clear input, there is ‘high Kalman gain’ (strong bottom-up influence). For the sake of the illustration, the probabilistic nature of the predictions and the interaction among predictions at different levels are not discussed. The implementation of the internal model in the language production system will not be discussed here. Figure reproduced from Pickering & Garrod (2007, p. 108) with permission © 2007 Elsevier

Figure 3. Schematic illustration of the phonological aspects of lexical prediction generated by a cerebellar internal model. Its organization is explicitly based on the well-studied cerebro-cerebellar circuitry and corresponding control-theoretic accounts of motor control (figure 1 above). A copy of a process involving the interaction of the prefrontal cortex with another brain area (such as the temporo-parietal cortex) in the context of comprehending a sentence is sent via the pontine nuclei to its interconnected cerebellar lobules. Cerebellar predictions are transmitted from the Purkinje cells of the cerebellar cortex via the ventrolateral neodentate nucleus and the thalamus back to cortical areas. Predicted and actual consequences of the process are compared in the inferior olive. Any discrepancies between the two are fed via climbing fibers from the inferior olive to the cerebellar cortex in the form of an error signal. LTD (long-term depression) is triggered at the parallel fiber-to-Purkinje cell synapses, updating the internal model. For the sake of the illustration, the cortico-cortical interactions are assumed to proceed in a bottom-up fashion (see section 7 for more discussion). See also Ito(2008) for discussion of available evidence from anatomical connectivity studies supporting these models. Same

simplifications are made as in Pickering and Garrod (2007; figure 2 here).

Table Legends

Table 1. Concreteness, imageability, and meaningfulness ratings from Coltheart (1981) and Brysbaert, Warriner, and Kuperman (2014) for the available verbs used in the ‘non-predictive’ and ‘predictive’ sentences in Lesage et al. (2012). Verbs in ‘predictive’ sentences are more concrete, imageable, and meaningful as compared to those in ‘non-predictive’ sentences. Stimuli retrieved from Lesage (2013).

Table 2. Peak activation coordinates (MNI) in posterolateral regions of HVIIa Crus I/II for: the contrast ‘Congruent > Scrambled’ in Moberget et al. (2014); the functional connectivity study of Krienen and Buckner (2009); contrasts in a selection of fMRI studies on default-mode network. The coordinates of Rao, Motes, and Rypma (2014) are converted from Talairach to MNI space.

MARKED REVISION

The Cerebellum, Internal Models and Prediction in 'Non-motor' Aspects of Language: A Critical Review

Georgios P. D. Argyropoulos

Developmental Neurosciences Programme
UCL Institute of Child Health
30 Guilford Street, London
WC1N 1EH
UK

Correspondence:

Dr. Georgios P. D. Argyropoulos (g.argyropoulos@ucl.ac.uk)

Abstract: The emergence of studies on cerebellar contributions in ‘non-motor’ aspects of predictive language processing has long been awaited by researchers investigating the neural foundations of language and cognition. Despite i) progress in research implicating the cerebellum in language processing, ii) the widely-accepted nature of the uniform, multi-modal computation that the cerebellum implements in the form of internal models, as well as iii) the long tradition of psycholinguistic studies addressing prediction mechanisms, research directly addressing cerebellar contributions to ‘non-motor’ predictive language processing has only surfaced in the last five years. This paper provides the first review of this novel field, along with a critical assessment of the studies conducted so far. While encouraging, the evidence for cerebellar involvement in ‘non-motor’ aspects of predictive language processing remains inconclusive under further scrutiny. Future directions are finally discussed with respect to outstanding questions in this novel field of research.

Keywords: language, prediction, cerebellum, internal models, associative learning, priming, default-mode network

1. *Introduction*

The cerebellum was traditionally seen as exclusively supporting the coordination of skilled, voluntary movement, gait, posture, balance, control of muscle tone, motor learning and articulation, with such involvement being reported for two centuries now. Readers are referred to Dow and Moruzzi (1958) for a thorough review of 19th century studies in the field, and to Manto et al. (2012) for a recent account. Over the last few decades, however, especially after the seminal work by the Leiners (see Leiner (2010) for a brief review), cumulative evidence has supported cerebellar involvement in a wide range of higher cognitive functions, e.g. memory, executive functions, visuospatial processing, emotional regulation, thought modulation, and, crucially, language (Mariën et al., 2014). In parallel, it has become increasingly clear that the cerebellum communicates in segregated anatomical loops with motor and prefrontal cortex (Strick, Dum & Fiez, 2009). Its cytoarchitectural uniformity supports the idea that its computations for motor control should guide hypotheses about its contributions in higher cognitive processes (Ramnani, 2006), including language (Ito, 2000a, 2008). However, research of relevance has, until very recently, been predominantly conducted in three poorly interfacing fields: i) work on cerebellar internal models in motor and non-motor aspects of behavior, without addressing their contribution in language; ii) research on cerebellar involvement in non-motor aspects of language processing, with no computational grounding; iii) studies conceptualizing predictive operations in language processing in terms of outputs of internal models, without addressing cerebellar circuitry as the most likely candidate for their implementation. This review will focus on studies of cerebellar involvement in the generation of semantic and phonological predictions above the lexical level in language comprehension. The interaction of predictive processes at different levels and the covert employment of language production mechanisms in comprehension (Pickering & Garrod, 2007) remain beyond the scope of this review.

2. *'Cerebellum and 'Prediction', but no 'Language'*

The contributions of the cerebellum in the automation of motor control have been well studied, and are consistent with the view that learning mechanisms store 'motor memory' through the diminishing synaptic efficacy between cerebellar neurons and their inputs (Marr, 1969; Albus, 1971). Control theoretic accounts suggest that such mechanisms allow cerebellar circuitry to acquire internal models that ultimately implement in an automatic fashion the movement-related processes initially established in the motor cortex (Ito, 1970, 1984; Wolpert, Miall, & Kawato, 1998; Ramnani, 2006). For instance, an internal (‘forward’) model of the arm’s dynamics, receives, as input, information on the current position and velocity of the arm, along with an ‘efference copy‘ of motor commands issued by the central nervous system, and outputs a prediction of the future position and velocity of the arm. Because of conduction delays in efferent and afferent pathways, the central nervous system is not immediately updated on changes in the peripheral motor system, and any recent commands issued may be yet to affect the musculature. Internal models are employed more rapidly, providing information about future properties of the controlled object, a fortiori in cases where accurate sensory feedback may be totally absent. This internal model ‘feedback’ allows the perceiver to rapidly interpret the perceptual signal and react accordingly, complete percepts received incompletely and/or under noise, and disambiguate in situations of uncertainty (Jordan & Wolpert, 2000). While internal models are conceivably located in all brain regions with synaptic plasticity that receive and send relevant information for their input and output (Kawato, 1999), a broad range of electrophysiological (Gilbert & Thach, 1977; Medina & Lisberger, 2009), imaging (Imamizu, Miyauchi, Tamada, Sasaki, Takino, Pütz, Yoshioka, & Kawato, 2000; Puttemans, Wenderoth, & Swinnen, 2005; Blakemore, Frith, & Wolpert, 2001; Desmurget, Gréa, Grethe, Prablanc, Alexander, & Grafton, 2001), and clinical studies (Müller & Dichgans, 1994; Nowak, Timmann, & Hermsdorfer, 2007) have established the cerebellum as ‘the most likely site for forward models to be stored’ (Kawato, Kuroda, Imamizu, Nakano, Miyauchi, & Yoshioka, 2003, p. 171).

However, the cerebellum exhibits a combination of two striking properties that support the involvement of its internal models beyond motor control: namely, its ‘essentially uniform, monotonously repetitive architecture’ (Schmahmann, 2000, p. 206) and its massive connectivity with virtually all major subdivisions of the brain. In particular, axonal fiber-tracing studies have identified projections from a broad range of neocortical areas to the ponto-cerebellar system, and even further to specific cerebellar lobules. These lobules return projections to the very same cerebral cortical areas via the cerebellar nuclei and thalamus, hence operating by means of segregated anatomical loops. Characteristically, the primary motor cortex selectively communicates with cerebellar lobules HV, HVI, HVIIb, and HVIIIa (Middleton & Strick, 2000; Kelly & Strick, 2003). Importantly, though, a substantial range of prefrontal areas send inputs to the pontine nuclei, which, along with the inferior olive, form the two major sources of input to the cerebellum. These prefrontal areas span from area 10 through to posterior regions of area 8, crucially including area 45B in the rostral bank of the inferior limb of the arcuate sulcus (Schmahmann & Pandya, 1997). Prefrontal neurons provide inputs to Purkinje cells in lobule VIIa and Crura I and II in lobule HVIIa (Kelly & Strick, 2003). This has encouraged the idea that '[i]f closed-loop circuits reflect a general rule, then all of the areas of cerebral cortex that project to the cerebellum are the target of cerebellar output' (Strick, Dum, & Fiez, 2009, p. 422). Furthermore, lobule HVIIa boasts unparalleled computational power: it comprises nearly 50% of the cerebellar volume (Balsters et al., 2010), while the human cerebellum itself contains more neurons than all the rest of the nervous system (see Leiner (2010) for references). Moreover, recent work employing intrinsic functional connectivity in humans has demonstrated that HVIIa Crus I/II can be further subdivided on the basis of functional connectivity with anterior, ventromedial, and dorsolateral regions in the prefrontal cortex (Krienen & Buckner, 2009; Buckner, Krienen, Castellanos, Diaz, & Yeo, 2011). Since the uniformity of cellular organization across the cerebellar cortex implies identity in the computations performed (e.g. Bloedel, 1992; Schmahmann, 1997), the same forms of plasticity might support the automation of cognitive processes, and knowledge on cerebellar motor learning has indeed started to guide the study of cerebellar contributions to cognition (Ramnani, 2006; figure 1 below).

Associative learning represents a fundamental mechanism by which cerebellar internal models are seen to work. Classical conditioning, for instance, provides the most basic form of associative memory formation, and the cerebellum has long been established as a fundamental site (e.g. Thompson, Bao, Chen, Cipriano, Grethe, Kim, Thompson, Tracy, Weninger, & Krupa*,* 1997; Christian & Thompson, 2005) in both comparative (e.g. McCormick & Thompson, 1984) and human clinical studies (e.g. Daum, Ackermann, Schugens, Reimold, Dichgans, & Birbaumer, 1993; Timmann, Kolb, Baier, Rijntjes, Mueller, Diener, & Weiller, 1996). Crucially, the cerebellum is involved in non-motor aspects of associative learning. In a series of studies, cerebellar patients acquired associations between colors and numerals by trial-and-error. In comparison with normal controls, patients were significantly slower in learning the correct associations, and were impaired in recognizing them later. Control conditions established that this learning deficit could not be reduced to patients' motor impairment (Drepper, Timmann, Kolb, & Diener, 1999; Timmann, Drepper, Maschke, Kolb, Böring, Thilmann, & Diener, 2002; Timmann, Drepper, Calabrese, Bürgerhoff, Maschke, Kolb, Daum, & Diener, 2004). Moreover, recent fMRI work has shown that HVIIa Crura I/II are involved in the acquisition and employment of first- and second-order rules. In these studies, subjects acquired arbitrary associations by trial-and-error of visual stimuli with manual responses (first-order rules) or with instructions on the selection of a first-order rule (second-order rules). With the employment of delayed response tasks, activity time-locked to the onset of the rule-related cue was disambiguated from that for motor responses or visual feedback (Balsters & Ramnani, 2008, 2011; Balsters, Whelan, Robertson, & Ramnani, 2013). These findings corroborate early proposals that the cerebellum 'predicts' and 'prepares' the internal conditions required for sensory, motor, autonomic, memory-related, attention-related, affective, or linguistic operations, by acquiring the 'predictive relationships among temporally ordered multidimensional sequences of exogenously derived […] and endogenously derived […] neural activities' (Courchesne & Allen, 1997, p. 2). More recently, contributions of cerebellar internal models to multi-modal associative learning have been thoroughly examined in two reviews (Timmann, Drepper, Frings, Maschke, Richter, Gerwig, & Kolb, 2010; Bellebaum & Daum, 2011). Quite importantly, though, neither of those discusses the ways in which cerebellar internal models could contribute to non-motor aspects of linguistic associative learning and processing.

[Please insert figure 1 here]

3. *'Cerebellum and Language', but no 'Prediction'*

In parallel, studies on the cerebellar contributions in language have started to flourish over the last few decades. Although still often marginalized in neurobiological models of language processing (e.g. Hickok & Poeppel, 2007; Friederici, 2012), the 'ongoing enigma' of the linguistic cerebellum has been attracting significant attention (Mariën et al., 2014). Indeed, there has been cumulating clinical evidence for even a ‘lateralized linguistic cerebellum’ (Mariën, Engelborghs, Fabbro, & De Deyn, 2001), and meta-analyses of imaging studies locate activations related to language processing in right hemispheric cerebellar regions (E, Chen, Ho, & Desmond, 2014; Stoodley & Schmahmann, 2009). For an up-to-date account of the evidence, the reader is encouraged to consult De Smet, Paquier, Verhoeven, and Mariën (2013), as well as papers in this special issue. Even so, such discussion has not incorporated the way in which cerebellar internal models may be involved in language processing, until very recently.

4. *‘Language and Prediction’, but no ‘Cerebellum’*

Turning to prediction mechanisms in language processing, these are anything but novel a concept in studies of psycholinguistics and the neurobiology of language. For decades now, research has shown that predictions at the sentence level modulate speech perception and production (Lieberman, 1963) and accelerate syntactic processing (e.g. Staub & Clifton, 2006; see also discussion in Sturt & Lombardo, 2005).

Unsurprisingly, then, prediction in language processing is often conceptualized to operate in the form of internal model outputs. These are used in work on auditory feedback for speech production (Perkell, Matthies, Lane, Guenther, Wilhelms-Tricarico, Wozniak, & Guiod, 1997; Golfinopoulos, Tourville, Bohland, Ghosh, Nieto-Castanon, & Guenther, 2011; Golfinopoulos, Tourville, & Guenther, 2010), providing rapid information on the predicted auditory consequences of articulatory gestures well before the later-arriving sensory feedback (Heinks-Maldonado, Nagarajan, & Houde, 2006; Tourville, Reilly, & Guenther, 2008), or even in the absence of overt articulation (Tian & Poeppel, 2010). The N400, an event-related brain potential response, has also been discussed within the context of violated predictions generated by internal models (Hosemann, Herrmann, Steinbach, Bornkessel-Schlesewsky, & Schlesewsky, 2013).

In their seminal paper, Pickering and Garrod (2007) advanced the idea that internal models are employed in language processing in a multi-level fashion, extending beyond acoustic perception. Language perception and comprehension were proposed to dynamically combine the outputs generated by an ‘input analysis system’ with predictions of internal models providing an estimate of the next state that the input analysis system will enter (figure 2 below). Whenever the prediction is strong and the input noisy, the internal model exerts strong influence to the 'input analysis system'*.* In that way, the internal model ensures noise-resistant, rapid perception and comprehension of utterances. It should be noted that Pickering and Garrod's (2007) work draws from literature discussing the covert, imitative involvement of action production mechanisms in the efficient perception of actions performed by conspecifics (Wilson & Knoblich, 2005). Whether cerebellar internal models in language comprehension recruit language generation mechanisms remains an outstanding question (see section 7 below).

[Please insert figure 2 here]

5. Cerebellum, Prediction and Language

Given the well-established nature of cerebellar contributions in motor control (Manto et al., 2012), studies of speech production and perception have synthesized insight from these three strands of research (cerebellar involvement in predictive processing; engagement of the cerebellum in 'non-motor' aspects of language; prediction mechanisms in language processing). Recent proposals emphasize the importance of the cerebellum and the basal ganglia in detecting temporal regularities and generating predictions during speech processing- hence alleviating the workload of fronto-temporal speech processing networks (Kotz & Schwartze, 2010). Accounts of language production also suggest that cerebellar internal models operate at the articulatory level (Hickok, 2012). Superior paravermal regions, for instance, may encode feedforward programs for the production of syllables (Ghosh, Tourville, & Guenther, 2009). Stuttering has been discussed within the context of deficient cerebellar internal models in generating a motor prediction error (Golfinopoulos, Tourville, & Guenther, 2010). Similarly, right cerebellar lobule HVI may undertake the temporal processing of segmental properties and left HVI may process prosodic, melodic properties (Callan et al., 2007). More recent work implicates the cerebellum in perceptual enhancement of acoustic properties of the linguistic signal. An fMRI study employed a word recognition task including acoustic stimuli from severely distorted speech. Improvements in the perception of such stimuli modulated cerebellar activity in four distinct cerebellar regions, one crucially being the right HVIIa Crus I. Activation in this lobule functionally correlated with cerebral regions that encompassed portions of the left angular and left temporal gyri. The findings supported the idea that regions within the left temporal and parietal cortex and the right Crus I (potentially along with lobules V/VI), participate in a functional network for achieving adaptive plasticity in speech perception. Discrepancies between the actual distorted acoustic speech input and the predicted acoustic input for a lexical item were discussed as engaging cerebellar-dependent supervised learning mechanisms (Guediche, Holt, Laurent, Lim, & Fiez, 2014).

The evidence accumulated so far has already encouraged discussions on the extension of the contributions of cerebellar internal models to non-motor aspects of language processing (Ito, 2000a, 2008; Argyropoulos, 2008, 2009). For instance, Argyropoulos (2009) argued that the cerebellum can store associative memory traces of contiguous linguistic events in sentence comprehension, such as the instance of processing the phonological, syntactic, and semantic properties of word ‘x’ and subsequently those of word ‘y’. In brief, cerebellar internal models would receive a copy of the input processed at each stage by a fronto-temporal ‘input analysis system’ (Pickering & Garrod, 2007) along with the operations performed on it, via the cerebro-ponto-cerebellar pathway. The internal models would then output a prediction of the next stage that the system would enter, via the ventrolateral neodentate, the parvocellular red nucleus and the thalamus back to the prefrontal cortex. Any discrepancies between the two would be conveyed as error signals back to the neocerebellar cortex through climbing fibers from the inferior olive. By long-term depression of parallel fiber-to-Purkinje cell synapses, these errors would train the internal model for accurate predictions. In cases where input analysis is conducted in noisy conditions and/or strong neocerebellar predictions are transmitted, neocerebellar output would override and/or bypass the output of cortico-cortical processing. With its massive computational power, lobule HVIIa may successfully undertake such an enormous task. Figure 3 provides a schematic illustration of phonological predictions generated in cerebro-cerebellar circuitry-conceivably, though, cerebellar circuitry would undertake the acquisition of temporal regularities and the implementation of predictions at different levels (e.g. phonology, syntax, semantics). While the specialization of distinct lobular regions in different levels of prediction remains an outstanding question in the field (see section 7 below), a series of studies have followed these proposals, providing evidence for the possibility of cerebellar internal models to transmit predictions on upcoming 'non-motor' information in sentence processing.

[Please insert figure 3 here]

5.1. Cerebellar involvement in phrasal predictions

In the first cerebellar TMS study to use a language task, Argyropoulos (2011a) addressed cerebellar contributions to such 'non-motor' aspects of predictive language processing. At a minimal level, the capacity of the neocerebellum to bias and/or pre-empt the prefrontal cortex for upcoming linguistic input was predicted to manifest itself in phrasal associative, and not semantic categorical lexical priming. On the one hand, phrasal associative priming pertains to the probability that one word may call to mind a second one, and is based on the temporal contiguity and predictability of items in discourse. Pairs of lexical morphemes in idiomatic phrases provide a case in point, such as 'gift-horse', 'skeletons-closet'. On the other hand, semantic categorical relatedness reflects the taxonomic relations between paradigmatic co-exemplars and the overlap in featural descriptions of two words, e.g. 'tree'-'bush', 'bee'-'grasshopper' (Thompson-Schill, Kurtz, & Gabrieli, 1998; Hutchison, 2003).

A lexical decision task assessed noun-to-noun priming sizes (quantified as differences in milliseconds between decision latencies for related and unrelated lexical pairs) before and after rTMS of the right neocerebellar vermis and of a control site in healthy native English speakers, as yielded by two different types of pairs: phrasal associates (e.g. 'gift'-'horse'), and pairs of subordinate and superordinate terms of the same category (e.g. 'penny'-'coin'). TMS of the neocerebellar vermis selectively enhanced phrasal associative priming in the form of increased differences between response latencies for lexical targets preceded by an unrelated prime (e.g. 'battery'-'horse') and those preceded by a phrasally associated one (e.g. 'gift'-'horse'). No such effects occurred for categorically related items after TMS, and stimulation of the control site did not affect associative priming sizes. This finding was consistent with evidence that cerebellar damage impairs verbal fluency by affecting phonemic rule-based word production, yet sparing semantic rule-based performance (Leggio, Silveri, Petrosini, & Molinari, 2000). Encouragingly, research in Alzheimer's disease and semantic dementia has established that temporal lobe lesions induce disruptions in categorical semantic, but not in associative priming (Rogers & Friedman, 2008). Moreover, cerebellar patients have been shown to ably perform category learning tasks, unlike patients with Parkinson's disease (Maddox, Aparicio, Marchant, & Ivry, 2005). This finding also echoes speculations on the significance of the cerebellum in storing and generating prefabricated, rote-memorized, idiomatic sequences, as opposed to propositionally composed ones: '[w]e learn and can recite ''Jabberwocky'' as movement and not at all as language […] We rote- memorize something that has so little linguistic or logical connection among the elements that it is learned as a movement. We can listen to what we say in order to get at what we otherwise can't remember […] But it is not something we know. It is buried in a rote-learning movement sequence' (Thach, 1997, pp. 91-2).

5.2. Cerebellar involvement in semantic predictions

Arguably, though, word-to-word predictions need not occur exclusively at the phonological level. There is substantial evidence in the lexical priming literature for semantic associations reflecting the predictive linkage of concepts based on world knowledge, as in 'instrument-action' pairs ('broom'-'sweep'), 'script relations' ('theatre'-'play'), 'locative relations' ('beach'-'house'), 'compositional relations' ('brick'-'house'; Hutchison, 2003; Jones & Estes, 2012). According to the theory on event schemata, stereotypical knowledge about the world is organized into units, including information about sequences of objects, situations, events, or states. Schemata are not the products of inferential manipulation of declaratively encoded representations. They operate online and below the level of awareness in sentence comprehension upon encountering the relevant concepts (e.g. the association of a dog with the agent-role in a biting event). In both dual-stage and constraint-based models of sentence comprehension, mechanisms of prediction (McRae, Hare, Elman, & Ferretti, 2005) and schema transmission (Ferreira, 2003; Townsend & Bever, 2001) employ these semantic associations in order to assign probable thematic (θ)-roles (Chomsky, 1981) and form a meaning-form hypothesis.

Suggestively, there is ample evidence from imaging, neurostimulation, and clinical studies in support of the involvement of the cerebellum in word generation tasks, crucially verb-to-noun generation. In a seminal PET study, subjects read aloud, repeated, passively read, or listened to nouns, or generated semantically appropriate verbs in response (e.g. ‘eat’ for ‘cake’). Blood flow changes occurred in paravermal regions when subjects read aloud or repeated nouns, as compared to when they viewed or listened to them. However, activation in the right lateral cerebellum was found for overt verb generation, but not for noun repetition/reading. This was difficult to account for on a motor basis. It was equally enigmatic though with respect to the involvement of the cerebellum in computations underlying verb generation (Petersen, Fox, Posner, Mintun, & Raichle, 1989). Recent fMRI studies provide similar results. Comparing verb generation and verb reading, Frings, Dimitrova, Schorn, Elles, Hein-Kropp, Gizewski, Diener and Timmann (2006) found activations in the right cerebellar lobule HVI and HVIIa Crus I as a measure of verb generation; again, these were lateral from the paravermal activation of lobule VI, which was associated with speech articulation. Fiez, Petersen, Cheney, and Raichle (1992) examined an English-speaking patient with a large right cerebellar infarct, who generated inappropriate responses in a number of word-generation tasks, involving verb generation. The patient’s responses were inappropriate, but remained categorically related (e.g. ‘small’, instead of take or ‘swallow’, in response to ‘pill’). This could not be attributed to overall cognitive impairment, as the patient’s performance on tests of memory, intelligence, ‘frontal function’, and language skills was excellent. This supports the idea that cerebellar damage leaves semantic networks intact (Fiez & Raichle, 1997). In another study, patients performed poorly in generating appropriate verbs for nouns, but selected the correct verb for a noun from a list of alternative responses, suggesting that semantic/syntactic representations were preserved. They were also able to produce appropriate subordinate term-responses to superordinate terms, suggesting that ‘[t]he right posterolateral cerebellum may be more involved in associative semantics than in categorical semantics’ (Gebhart, Petersen, & Thach, 2002). This pattern of impairments is qualitatively different from the direct disruptions in semantic networks associated with temporal cortical pathology in semantic dementia and Alzheimer’s disease (e.g. Rogers & Friedman, 2008). Studies on ‘associative’ and ‘semantic’ priming in Broca’s and Wernicke’s aphasia also show no impairments in automatically accessing lexical and/or semantic information (Blumstein, Milberg, & Shrier, 1982; Milberg & Blumstein, 1981), suggesting that associative priming does not rely on these perisylvian structures. As already discussed, cerebellar patients ably perform category learning tasks, unlike patients with Parkinson's disease (Maddox et al., 2005). Moreover, in a study employing a card-sequencing task, cerebellar patients with left lesions showed selective impairments in script sequences based on pictorial material, while those with right lesions were only impaired in script sequences requiring verbal elaboration (Leggio, Tedesco, Chiricozzi, Clausi, Orsini, & Molinari, 2008). Likewise, cerebellar cathodal tDCS has facilitated the rate and consistency of responses in a verb generation task, as compared to sham and anodal stimulation. These facilitatory effects occur in the case of cerebral cortical tDCS following anodal stimulation. They were hence explained in terms of disinhibition of the left prefrontal cerebral cortex resulting from the inhibitory effect of cathodal tDCS on the overall inhibitory tone that the cerebellum exerts on the cerebral cortex (Pope & Miall, 2012). Finally, in a TMS study, right cerebellar rTMS reduced category switching in the form of reduced phonemic and semantic fluency (Arasanz, Staines, Roy, & Schweizer, 2012).

The involvement of neocerebellar circuitry in the generation of semantic predictions was first directly assessed in a cerebellar rTMS study (Argyropoulos, 2011b; Argyropoulos & Muggleton, 2013). A pair-wise lexical decision task used noun-primes, the semantic properties of which could be categorically related by synonymy (e.g. ‘theft’-‘stealing’) or thematically associated (denoting agents, patients, instruments or locations of actions) with their verb-targets (e.g. ‘chef’-‘cooking’). Four groups of subjects were employed: a group that underwent stimulation of a medial cerebellar site; a group that underwent stimulation of a right lateral cerebellar site corresponding to a region in right HVIIa Crus I involved in cognitive aspects of verb generation (Frings et al., 2006); two groups that completed the same session with no intervention of TMS, one with a 5-minute break and another one without. Stimulation of this lateral cerebellar site selectively boosted associative priming. No effects appeared after medial cerebellar stimulation or no stimulation. Argyropoulos (2011b) and Argyropoulos and Muggleton (2013) discussed this effect as a disruption of inhibitory processes in generating predictions for associated events expressed by the verb-target, e.g. (cooking)´, upon processing the semantic properties of the noun-prime, e.g. (chef)´.

This study was soon followed up by Lesage, Morgan, Olson, Meyer, and Miall (2012; see also Lesage (2013) for stimuli used). The authors employed the ‘Visual World’ paradigm (Altmann & Kamide, 1999), assessing the onset latencies of listeners’ saccadic eye movements towards images of objects that aurally presented sentences referenced. The authors used two types of sentences: ‘predictive’ sentences, where the object could be predicted among four alternatives (one target object and three distractor objects) on the basis of the semantic content of the verb (e.g. ‘The man will sail the … boat/mountain/bird/car’), and control, ‘non-predictive’ sentences (e.g. ‘The man will watch the … boat/mountain/bird/car’), where the verb was not selective for the target object. In pre-TMS conditions, priming occurred in the form of faster anticipatory responses in the predictive as compared to the non-predictive condition. Cerebellar rTMS significantly delayed such anticipatory responses. Stimulation did not change saccadic latencies in the non-predictive sentences, ruling out a general effect on language processing. It did not change eye movement kinematics either, thus ruling out disruptions in oculomotor control. The prediction deficit was moreover absent in two control groups (vertex stimulation and no stimulation), ruling out non-specific effects of stimulation.

More recently, two fMRI studies have attempted to address this issue. Both used sentence comprehension tasks manipulating cloze probabilities and hence the strength of ‘semantic priming’ (Moberget, Gullesen, Andersson, Ivry, & Endestad, 2014; Lesage, Hansen, & Miall, 2014; unpublished). The task of the first study involved the presentation of a sequence of five centrally presented words. The authors manipulated cloze probability by varying the context provided by the sequence of the initial four words. In a ‘Congruent’ condition, the target word was highly predictable on the grounds of the preceding four (e.g., ‘two plus two is four’). In an ‘Incongruent’ condition, a final word was also highly predictable, but that prediction was violated by the target word presented (e.g., ‘[the water] had frozen to cars’). In a ‘Scrambled’ condition, the first four words made a non-grammatical sentence stem (e.g., ‘fast in clock plane’), rendering the target word unpredictable (e.g., ‘through’). In the end of each presentation, participants assessed whether the sentence was meaningful or not by a button press. In their contrasts of interest, ‘Congruent > Scrambled’ yielded a cluster of activation across right HVIIa Crus I/II.

'Incongruent > Congruent' yielded an activation cluster across left HVIIa Crus I/II, IV, and medial VIIa Crus I and another across right HVIIa Crus I/II, medial VIIa Crus I/II and VI. 'Incongruent > Scrambled' showed activation in right HVIIa Crus I/II, VIIa Crus I/II, and VI, and left HVIIa Crus I/II and VIIa Crus I. The authors reasoned that both the generation of a prediction for an upcoming word based on its sentential context as well as the violation of that prediction are consistent with the involvement of neocerebellar internal models in non-motor aspects of language processing (Moberget et al., 2014).

Using an event-related fMRI design, the authors of the second study demonstrated that sentence predictability modulated the BOLD signal amplitude in a right HVIIa. Three events were modelled per trial: a context sentence (e.g. 'Greg went home for Christmas dinner'); the stem of a second sentence (e.g. 'His mum always cooked a'); the end of the second sentence (e.g. 'turkey'). Subjects read the sentences and pressed a button indicating the plausibility of the outcome. Crucially, the stem event did not require a motor response. Predictability was used as a parametric modulator for context and stem events. The authors showed that 'a cluster in the right posterolateral cerebellum […] was modulated by the predictability of the stem independent of outcome' (Lesage, Hansen, & Miall, 2014; see Lesage (2012, pp. 194-5) for stimuli).

5.3. Cerebellar contributions to the routinization of language processing

Cerebellar internal models have been held not only to generate predictions about the next state of a simulated process, but also to gradually undertake its automatic implementation (Ito, 1984; 2008). Interestingly, casual adult dialogical interaction exhibits phenomena of routinized language processing, with lexical items and syntactic constructions of re-occurring phonological and semantic properties (Pickering & Garrod, 2004). The significance of the cerebellum in automatizing motor repertoires has been supported by clinical (e.g. Lang & Bastian, 2002) and imaging evidence (e.g. Jueptner, Frith, Brooks, Frackowiak, & Passingham, 1997; Toni, Krams, Turner, & Passingham, 1998), extending recently to higher cognitive aspects of rule learning (e.g. Balsters & Ramnani, 2011). In a tDCS study, cerebellar stimulation impaired the practice-dependent improvement in performance in the Sternberg task irrespective of polarity. On the contrary, tDCS over the dorsolateral prefrontal cortex left it unchanged, markedly affecting verbal working memory per se.

Cerebellar stimulation also left visual evoked potentials unchanged, thus excluding visual cortex involvement (Ferrucci, Marceglia, Vergari, Cogiamanian, Mrakic-Sposta, Mameli, Zago, Barbieri, & Priori, 2008).

Crucially, cerebellar impairments induce similar automatization deficits in language-related processes. In word generation tasks, cerebellar patients reduce their response latencies poorly across repeated blocks of trials in comparison with normal controls (Fiez et al., 1992; Gebhart, Petersen, & Thach, 2002). Further evidence has been recently provided by Argyropoulos, Kimiskidis, and Papagiannopoulos (2011). The authors recruited two groups of participants of native speakers of Modern Greek for a TMS study conducted in two sessions, with the same stimuli presented in the second session. One group received stimulation centered at a site overlying the neocerebellar vermis in the first session, and on a deeper, control site in the second. Another group received stimulation in these two sites in reverse order. The study employed a pairwise lexical decision task. Lexical decision latencies for the group that underwent stimulation of the neocerebellar vermis in the second session did not become any shorter after stimulation, in contrast to all other conditions. Such disruption could not be explained on the grounds of sensorimotor processes, since the second group, who received neocerebellar vermal stimulation in their first session, showed a significant reduction of their reaction times in that session after stimulation. However, the disruption only appeared after stimulation of the target site selectively for the first group, who were encountering letter-string pairs for the second time. The automatization deficit thus pertained in particular to processing pairs that had been already encountered.

6. *Critical assessment*

The evidence presented here for the involvement of the cerebellum in ‘non-motor’ aspects of predictive language processing is consistent with the unitary, multi-modal nature of the contributions of cerebellar internal models. It appears also to follow from the sizeable volume of suggestive evidence that has preceded it. However, as this section will demonstrate, there are significant issues that compromise the strength of these findings and limit their interpretation.

6.1. *TMS studies*

Despite the significant advantages of cerebellar neurostimulation with TMS (Grimaldi et al., 2013) and tDCS (Grimaldi et al., 2014) over cerebellar patient and fMRI studies, there are methodological issues that are often hard to overcome. For instance, in TMS, optimal localization of the target region is meaningfully achieved by retrieving coordinates of the area with the highest level of activation during performance of the same task in a previously completed fMRI session by the very same subject. Alternatively, high-resolution structural images may be used to position the coil according to the subject's underlying anatomy. When these are unavailable, scalp-based measurements from clear external landmarks are often used, such as the vertex or the inion (Devlin & Watkins, 2007). Indeed, all four TMS studies discussed here (Argyropoulos, 2011a; Argyropoulos, Kimiskidis, & Papagiannopoulos, 2011; Argyropoulos & Muggleton, 2013; Lesage et al., 2012), used the inion as an external landmark, thus making precise localization of the lobular cortex stimulated impossible. However, all four studies demonstrate further weaknesses that interact with the aforementioned limitation and compromise the inferences that can be drawn.

6.1.1. Stimulating the neocerebellar vermis

The putative target in the study of Argyropoulos (2011a) and Argyropoulos, Kimiskidis, and Papagiannopoulos (2011) was a right superior posterior vermal site. Their external landmark coordinates (1 cm below the inion, 1 cm laterally to right) targeted the medial (VIIa) Crus I/II. In order to estimate the depth of the site and thus the possibility of its successful stimulation, the authors recruited a volunteer whose brain image was already registered with a TMS-MRI co-registration system. The site corresponded to the right VIIa. Encouragingly, the superior posterior vermis is one of the lobules closest to the TMS coil (Miall & Christensen, 2004), and its stimulation has induced behavioral effects with high spatial precision (Hashimoto & Ohtsuka, 1995).

However, the anatomical connectivity of the vermis with the frontal lobes remains unclear (e.g. Kelly & Strick (2003) on medial VIIa-BA 46 connectivity; but also Coffman, Dum, & Strick (2011) on vermal-motor cortical connectivity). On the other hand, resting-state functional connectivity studies demonstrate that the vermis and the medial regions of the Crura contain a broad range of network nodes, the most posteromedial of which belongs to the dorsal attention network (Buckner et al., 2011). In another study, a cluster spanning across Crus II with the medial portions close to the vermis showed connectivity with the left executive control network (Habas, Kamdar, Nguyen, Prater, Beckmann, Menon, & Greicius, 2009). More recently, Halko, Farzan, Eldaief, Schmahmann, and Pascual-Leone (2014) applied cerebellar rTMS guided by subject-specific connectivity to evaluate the relevance of connections between cerebral and cerebellar hubs belonging to different functional networks. One of their stimulation sites was lobule VII, and they demonstrated that its stimulation influences the cerebral dorsal attention system.

Equally enigmatic is the nature of vermal contributions in language. On the one hand, vermal lobule VII is involved in oculomotor control (e.g. Hashimoto & Ohtsuka, 1995) and also supports emotional processing (E, Chen, Ho, & Desmond, 2014; Stoodley & Schmahmann, 2009). On the other hand, superior posterior vermal activations are found in studies of language processing that are hard to reduce to motor effects (Desmond & Fiez, 1998; Desmond, Gabrieli, & Glover, 1998). Vermal tumor resection in children is associated with agrammatism (Riva & Giorgi, 2000). Voxel-based morphometric studies also report correlations between vermal grey matter and working memory measures (Ding, Qin, Jiang, Zhang, & Yu, 2012), and other studies find increased activity in the vermis after lexical training (Raboyeau, Marie, Balduyck, Gros, Démonet, & Cardebat, 2004). Vermal atrophy is also the most widely-cited cerebellar abnormality in schizophrenia (Picard, Amado, Mouchet-Mages, Olié, & Krebs, 2008): indeed, lexical priming studies on schizophrenic patients show greater priming for associated-only word pairs than for pairs related only semantically or both semantically and associatively (Nestor, Valdman, Niznikiewicz, Spencer, McCarley, & Shenton, 2006).

Evidently, though, the use of external landmarks for coil positioning could not exclude the probability that adjacent lobular cortices of VI, VIIb and VIII were also stimulated. While stimulation of right paravermal compartments is traditionally achieved by placing the coil 2 cm laterally to the right from the inion (e.g. Miall & Christensen, 2004) instead of 1cm, the scalp coordinates used here for targeting the right neocerebellar vermis may concomitantly stimulate portions of the right superior paravermis (e.g. Hashimoto & Ohtsuka, 1995). These lobules are reciprocally connected with the motor cortex in a segregated fashion (Kelly & Strick, 2003). While the second priming type used (semantically related and unrelated pairs) provided sufficient conditions for controlling for motor effects, the interpretation of the main effect here would be different. It can only be speculated that the selective effect on associative priming would pertain to cerebellar involvement in ‘silent/covert speech’ (e.g. Ackermann, Wildgruber, & Grodd, 1998). In an fMRI study, activation within the same region was found during silent recitation of the names of the months of the year (Ackermann, Mathiak, & Ivry, 2004), i.e. a routinized sequence of lexical morphemes. Such ‘inner speech’ has been held to provide a common platform for a broad range of cognitive functions implicating the cerebellum (Ackermann, Mathiak, & Riecker, 2007). This account would echo discussions whereby formulaic utterances are ‘buried in a rote-learning movement sequence’ engaging the cerebellum (Thach, 1997, pp. 91-2). In that way, it could only be speculated that the abnormalities in predictive functions were owed to impairments in the covert employment of the language production circuit (Pickering & Garrod, 2007).

For the study in Argyropoulos (2011a), in particular, a series of other weaknesses would include the low number of subjects and items per condition. Moreover, the within-subjects design involved participants being exposed to the same items twice (stimulation of control and target site counterbalanced across subjects). This may have introduced confounds with the effects that cerebellar stimulation has on practice-induced facilitation in performance (Ferrucci et al., 2008; Argyropoulos, Kimiskidis, & Papagiannopoulos, 2011). Also, the study in Argyropoulos, Kimiskidis and Papagiannopoulos (2011) failed to replicate the selective effects of neocerebellar vermal stimulation on associative priming. This could be attributed to the unavailability of published word association norms in Modern Greek for the construction of the stimulus set. On the other hand, the main finding in Argyropoulos, Kimiskidis, and Papagiannopoulos (2011) on the effects of neocerebellar vermal stimulation on the disruption of practice-induced accelerations of lexical decisions when participants encountered the same stimuli for the second time was not present in Argyropoulos (2011a). This could be due to the higher number of participants employed in that study.

6.1.2. Stimulating the neocerebellar hemispheres

Similarly, in Argyropoulos & Muggleton (2013), the target stimulation site was identified by external landmarks. The aim was to stimulate the peak coordinates of the activation cluster yielded in Frings et al. (2006) when comparing verb generation in inner speech with verb reading in inner speech (right HVIIa Crus I). Its coordinates were again registered and converted into scalp coordinates using a TMS-MRI co-registration system in one volunteer, and corresponded to 10 cm laterally to the right from the inion. The significant distance from that landmark would only increase the inter-subjective variability and the uncertainty on the cerebellar lobule stimulated. Moreover, the significant depth of the underlying tissue may have made it impossible to successfully induce an effect in at least a subset of participants with the figure-of-eight coil used (Hardwick, Lesage, & Miall, 2014), despite the high amplitude of stimulation. Another concern is that significant associative priming was only observed in the second phase of the experimental session across the four groups (control TMS, target TMS, no TMS 1, no TMS 2). While the associative boost after stimulation of this lateral site was significantly larger than that for the other groups and was also absent from the semantic categorical priming type, a concrete interpretation of the effect remains elusive. It would suggest that the processes affected were not automatic in nature and may reflect TMS effects on strategically employing semantic associations.

Similar issues in identifying the exact cerebellar lobular cortex underlying the stimulation site apply in Lesage et al*.* (2012), given that external landmarks were used again. However, the authors successfully stimulated the deeper hemispheric regions of the neocerebellum by employing a double-cone coil (Hardwick, Lesage, & Miall, 2014). As compared to figure-of-eight coils, though, focality of stimulation is more limited, and the possibility of a spread of the effect in neighbouring areas is increased. Suggestively, the authors used a set of coordinates (1 cm below the inion and 3 cm laterally to the right) that have been traditionally employed to induce cerebellar inhibition and in return facilitation in MEPs evoked by single pulse TMS of the contralateral primary motor cortex (e.g. Oliveri, Koch, Torriero, & Caltagirone, 2005). These coordinates have also been used by the same laboratory in recent assessments of successfully modulating motor cortical excitability with cerebellar stimulation by different coil types (Hardwick, Lesage, & Miall, 2014). They have also been employed in their previous TMS investigations to aim at the hand area of the ipsilateral cerebellar cortex, with its stimulation affecting cortico-cerebellar projections, changing motor cortical excitatability in the contralateral hand area (Miall, Christensen, Cain, & Stanley, 2007). Previous studies have demonstrated that this target site ‘affects the hand area of motor cortex, consistent with activation of lateral cerebellar cortex, probably in lobules V and VI’ (Miall & King, 2007, p. 576). Therefore, despite the fact that in Lesage et al*.* (2012) stimulation with the same coordinates was ‘directed towards Crus II’ (Grimaldi et al., 2013, p. 133), motor-projecting cerebellar lobules HV and HVI were clearly also stimulated.

However, this limitation further interacts with a weakness in the construction of the stimulus set of the task. The authors acknowledge that they did not distinguish action-related from non-action related verbs, and were thus unable to test for any specific effects of action verbs in cerebellar motor-projecting lobules. They entertain the possibility that these are more active when processing action-related verbs. They reason, however, that, if such mechanisms were indeed present, they would not suffice to explain their results, given that the impairment was specific to sentences with predictive verbs, irrespective of any action-related semantics. The certainty with which such an interpretation can be dismissed comes into question when the stimuli (Lesage, 2013, p. 189) used in this study (Lesage et al., 2012) are examined under further scrutiny. The verbs in the ‘predictive’ condition score higher in concreteness, imageability, and meaningfulness metrics (e.g. Coltheart, 1981; Brysbaert, Warriner, & Kuperman, 2014) than those in the ‘non-predictive’ (control) condition. This is illustrated in table 1 below. The cerebellar lobules HV and HVI that were concomitantly stimulated form a segregated loop with the motor cortex (Kelly & Strick, 2003), and ample evidence associates action-related semantic processing with activity in motor structures used to implement those actions. For instance, processing speed for leg-related words is affected by TMS over the leg area of the left primary motor cortex, while performance remains unaffected when the left arm area or the right leg area is stimulated (Pulvermüller, Hauk, Nikulin, & Ilmoniemi, 2005). Similarly, fMRI studies show activations in passive reading of verbs that denote actions executed by different effectors are somatotopically organised in the primary motor cortex (e.g. Hauk, Johnsrude, & Pulvermüller, 2004). Another study has demonstrated that the MEP size in each effector muscle is only affected when listening to sentences containing actions related to that effector (Buccino, Riggio, Melli, Binkofski, Gallese, & Rizzolatti, 2005). Moreover, cathodal tDCS on the left motor cortex in healthy participants has reduced success rates in the acquisition of action-related words, in comparison with anodal or sham stimulation, tDCS over the prefrontal cortex, and learning of object-related words (Liuzzi et al., 2010). Unsurprisingly, imagined movement also engages the cerebellum (e.g. Hanakawa et al., 2008). Some first suggestive evidence for the modulation of processing action-related verbs by cerebellar TMS can be found in Oliveri et al. (2009).

[Please inset table 1 about here]

What if stimulation was to selectively affect HVIIa Crus I/II, which communicates with the contralateral dorsolateral prefrontal cortex (Kelly & Strick, 2003)? The confounds in the design of the stimuli between sentence ‘predictiveness’ on the one hand and verb imageability, concreteness, and meaningfulness on the other would still not support the conclusion that cerebellar TMS disrupted cerebellar predictive processing. This is because of the modulation of activity in the dorsolateral prefrontal cortex by such lexical semantic properties. Concrete as compared to abstract concepts elicit greater activity in the left dorsolateral prefrontal cortex (Roxbury, McMahon, & Copland, 2014). Stimulation of the left dorsolateral prefrontal cortex also preferentially affects verbs compared to nouns (Shapiro, Pascual-Leone, Mottaghy, Gangitano, & Caramazza, 2001; Cappa, Sandrini, Rossini, Sosta, & Miniussi, 2002). This supports the interpretation that the functional link with motor cortical hand or leg regions may be mediated by the dorsolateral prefrontal cortex, much as the ventrolateral prefrontal cortex mediates the connection between speech perception and the motor cortical mouth region (Watkins & Paus, 2004): rTMS of the left dorsolateral prefrontal cortex has decreased naming latencies for verbs relative to its right homologue and sham stimulation, leaving latencies for object naming unaffected (Cappa et al., 2002). In another experiment, participants inflected nouns and verbs of a natural language as well as their pseudoword counterparts. Stimulation over the left dorsolateral prefrontal cortex affected reaction times selectively for verbs and not nouns (Shapiro et al., 2001). Therefore, the selective effect that cerebellar TMS had on ‘predictive’ sentences may be explained by the fact that the properties of the verbs in these sentences engaged to a significantly larger extent action-related semantic processing in the motor and/or dorsolateral prefrontal cortex than the more abstract and less imageable and meaningful verbs in ‘non-predictive’ sentences did. Stimulation of motor- or prefrontal-projecting cerebellar lobules may have thus disrupted this engagement.

6.2. fMRI studies

Turning to the two fMRI studies recently reported (Moberget et al., 2014; Lesage et al., 2014), a concern in the design of both is the confound between processing difficulty and ‘predictability’. As it shall be argued, the activations in the particular regions of HVIIa Crus I/II that the authors report are probably yielded by differences in generic processing demands and not predictability per se.

6.2.1. Events: predictive or predicted?

A question for both fMRI studies would pertain to the nature of the events of interest. The activations in both studies relate to activity time-locked to the onset of a stimulus (word or sentence stem) that is constrained to a smaller or larger extent by the preceding context. It is therefore not clear whether those activations reflect cerebellar involvement in generating predictions for upcoming language input or cerebellar modulation by the processing demands for input constrained to a smaller or larger extent by preceding context.

Moreover, in Moberget et al. (2014) the activity at the onset of the final word was not sufficiently disambiguated from and was susceptible to contamination by that for response preparation and implementation. Their sentence types did not only differ with respect to the predictability of the last word, but also with respect to the demands in response preparation. The activations they report in the contrast ‘Incongruent > Congruent’ provide a characteristic case in point. Indeed, the activations in this contrast fall within the region of lobule HVIIa Crus I/II that shows functional connectivity with Krienen and Buckner’s (2009) dorsolateral prefrontal cortical map and the frontoparietal control network (Buckner et al., 2011). However, this contrast need not reflect a violation of prediction in sentence comprehension. The difference in predictability is confounded with a number of discrepancies in other demands. Incongruent sentences here invite increased attention to and inhibitory control of an action being prepared (button key press to signal positive response to the question of meaningfulness) but needs to be revised (different button key press to signal negative response) upon encountering the last word; elevated demands for checking morpho-syntactic feature consistency and achieving semantic integration; elevated verbal working memory demands in reanalysing the semantic and syntactic structure of the sentence stem upon encountering the unpredictable target word.

In Lesage et al. (2014), such confounds are avoided by sufficient jittering of the delay preceding the first context sentence and the second sentence stem. However, the content of the second sentence stem (the event of interest) reflects the predictions generated on the grounds of the first sentence context but also predicts the occurrence of the target word of the second sentence. The next section will make the case that these findings may reflect the modulation of hemodynamic activity by task difficulty, which is only confounded with predictability.

6.2.2. Networks: task-positive or task-negative?

Moberget et al*.* (2014) employed scrambled sentence stems as control contexts with minimal predictability for the final target word. In fact, the authors argued that, in studies similarly contrasting scrambled sentences with semantically and syntactically canonical sentences, 'the critical variable was the predictability of the presented words– with increased cerebellar involvement for predictable relative to unpredictable conditions.' (*ibid*, p. 2876). Yet scrambled sentences do not provide an unproblematic baseline if prediction per se is addressed in contrasting 'Congruent > Scrambled'. One major source of sentence comprehension difficulty is the interference of material that partially matches syntactic and semantic retrieval cues (see Glaser, Martin, Van Dyke, Hamilton, & Tan (2013) for discussion). Moreover, such anomalous sentences may engage the semantic integration process to a greater degree in participants' attempt to comprehend the sentence (Brown & Hagoort, 1993). For example, syntactic violations traditionally activate areas involved in syntactic processing, since structure building, agreement checking and other putative operations are disrupted, and hence extra attention is paid to these aspects (Kaan & Swaab, 2002). Greater activations often occur, for instance, in the left inferior frontal gyrus for violated sentences than for normal sentences (see discussion in Hagoort et al., 2009), and indeed scrambled sentences involve syntactic and semantic violations to a larger extent than sentences with no such violations. Cognitive control and conflict monitoring are arguably engaged more in scrambled sentences as compared to the other sentence types. Maintenance-related processes are also involved in keeping information accessible whenever the currently described events seize to be consistent with the global situation model under construction at each stage. With respect to sentence comprehension, then, scrambled sentences are arguably more taxing than congruent ones. This would mean that the ‘Congruent > Scrambled’ contrast would be reflecting not (just) ‘Predictable > Not predictable’, but also a more generic ‘Low Demands > High Demands’ discrepancy, owed to the larger demands in semantic and syntactic integration, cognitive control, conflict monitoring and more attempts for reanalysis made in the ‘Scrambled’ as compared to the ‘Congruent’ condition.

How then would a ‘Low Demands > High Demands’ contrast yield this activation in the posterolateral regions of HVIIa Crus I/II? Functional connectivity studies make clear that lobule (H)VIIa does not form a functionally unitary locus. While findings in different studies are not unequivocal, they show that there are distinct regions within this lobule that are functionally connected with different prefrontal cortical areas. In particular, they demonstrate that, apart from the well-established involvement of regions in (H)VIIa in the executive control network, there are distinct regions within this lobule that form crucial hubs of the default-mode network. These regions occupy a sizeable posterolateral portion of HVIIa Crus I/II (Krienen & Buckner, 2009; Buckner et al., 2011; Bernard et al., 2012; for discussion on the absence of such findings in Habas et al. (2009), see Buckner et al. (2011), p. 2324 and p. 2340). Hubs of the default-mode network consistently demonstrate increased activity during rest or low-demand tasks as opposed to high-demand tasks across a broad range of paradigms, and have been argued to support internal self-reflective thought. The suppression of this network in attention-demanding tasks is manifested in the form of ‘task-induced deactivations’, traditionally yielding significant activation in contrasts of ‘ Low Demands > High Demands’, given their larger negative BOLD signal amplitude in conditions posing higher demands (e.g. Mazoyer, Zago, Mellet, Bricogne, Etard, Houde, Crivello, Joliot, Petit, & Tzourio-Mazoyer, 2001; Binder, Frost, Hammeke, Bellgowan, Rao, & Cox, 1999; Shulman, Fiez, Corbetta, Buckner, Miezin, Raichle, & Petersen, 1997; McKiernan, Kaufman, Kucera-Thompson, & Binder, 2003; Gusnard & Raichle, 2001; Raichle, MacLeod, Snyder, Powers, Gusnard, & Shulman, 2001; Raichle & Gusnard, 2005).

There is indeed substantial evidence for the task-induced deactivation of these specific posterolateral regions of HVIIa Crus I/II, a fortiori in non-linguistic tasks and with no apparent demand for predictive processing. In one such study, participants were required to indicate whether a visually presented digit-symbol probe-pair was present or absent in an array of nine digit-symbol probe-pairs. Along with the deactivation of other major hubs of the default-mode network, the data also demonstrated strong deactivations in posterior regions of right lobule HVIIa Crus I/II (Rao, Motes, & Rypma, 2014). Elsewhere, Harrison et al. (2008) studied correlated activity fluctuations of the default mode network regions during three conditions: rest with eyes closed; a moral dilemma task; a Stroop task. They showed a striking uniformity in the anatomy of the default mode network across these conditions. Crucially, one of its hubs was a posterolateral region of HVIIa Crus I, showing activation in the resting state and moral dilemma, but deactivation in the Stroop task. Fransson (2005) also reports that one of the brain regions that correlated positively with other hubs of this network (precuneus / posterior cingulate cortex) during resting state and in a condition with eyes closed was the left and right posterolateral region of HVIIa Crus I. In another study (Fransson, 2006), one such node that showed significantly more activity during rest compared to a two-back working memory task was represented by a substantial cluster in posterolateral regions of the right HVIIa Crus I . Gusnard, Akbudak, Shulman, and Raichle (2001) examined hemodynamic activity related with judgments in an internally cued and an externally cued condition. Along with medial prefrontal cortical activations, the contrast of ‘Internally > Externally cued judgments’ yielded an activation in the posterolateral region of right HVIIa Crus I/II. Moreover, the HVIIa Crus I/II peak activation coordinates in these contrasts are markedly adjacent to those found in Krienen and Buckner (2009) as peak coordinates of functional connectivity with their medial prefrontal cortical map- a major hub of the default mode network. The involvement of these regions in the default-mode network need not negate their computational properties in implementing internal models. It may simply suggest that internal models in these regions support processes of the default-mode network. Table 2 illustrates the adjacency of the peak activation coordinates from Moberget et al. (2014; ‘Congruent > Scrambled’) with those reported in studies on the engagement and disengagement of the default-mode network. Highly consistent with this interpretation is also the activation that Moberget et al. (2014) report in lobule IX for both ‘Congruent > Scrambled’ and ‘Incongruent > Scrambled’, as lobule IX is also part of the default mode network (Habas et al., 2009).

[Please insert table 2 here]

Likewise, in Lesage, Hansen, and Miall (2014), the parametric modulation by cloze probability in a right posterolateral region of HVIIa Crus II may not exclusively pertain to predictability. Lower cloze probability entails higher attentional and verbal working memory demands, as the memory trace of preceding linguistic information needs to be refreshed in order to check against a range of different completions. Indeed, the magnitude of deactivation of the default-mode network increases in a graded fashion in accordance with task load and error-proneness (McKiernan et al., 2003; Singh & Fawcett, 2008; Polli, Barton, Cain, Thakkar, Rauch, & Manoach, 2005; Persson, Lustig, Nelson, & Reuter-Lorenz, 2007; Kincses, Johansen-Berg, Tomassini, Bosnell, Matthews, & Beckmann, 2008; Esposito, Bertolino, Scarabino, Latorre, Blasi, Popolizio, Tedeschi, Cirillo, Goebel, & Di Salle, 2006). Hence the modulation of activity in this posterolateral region of right HVIIa Crus II by predictability may reflect the modulation of the disengagement of the default mode network by task difficulty. This would be a by-product of predictive processing, which may instead be carried out by cerebral cortical regions.

6.3. *Directions for improvements*

The discussion here has made clear the need for a number of improvements in the methods and considerations on the different findings. For both TMS and fMRI studies, a substantial improvement would address the need for carefully designed stimulus sets, whereby distinct experimental conditions would be exclusively matched in all other aspects apart from those of interest, i.e. the predictability of upcoming stimuli. In any other case, modulation of cerebellar engagement by predictability may be confounded with other mechanisms which cerebellar circuitry may also support, such as those in processing action-related semantics (e.g. Lesage et al., 2012). Both TMS and fMRI studies should also take into account the functional connectivity of different regions within (H)VIIa Crus I/II with distinct prefrontal cortical areas. While a great part of this lobule is embedded in the executive control network, the existence of default-mode network hubs in certain posterolateral regions of HVIIa Crus I/II (e.g. Buckner et al., 2011; Bernard et al., 2012) suggests that these regions may show modulation by task difficulty as other hubs of this network in the cerebral cortex (e.g. Polli et al., 2007; Singh & Fawcett, 2008). This should be taken into account when interpreting the modulation of activity in lobule (H)VIIa by predictability in language processing, especially when activations occur within these posterolateral regions of lobule HVIIa (e.g. Moberget et al., 2014). Future studies would thus benefit from the use of further control conditions and/or tasks that concomitantly vary with respect to processing demands. These would help establish whether modulation of cerebellar activity by predictability is reduced to generic modulation by task load. Alternatively, future studies should show that linguistic predictability modulates activity in regions other than those that non-linguistic predictability does, or that different types of linguistic prediction (e.g. phrasal, semantic) yield activations in distinct cerebellar lobular regions. For TMS studies in particular, double-cone coils should be used to provide appreciable strength in the stimulation of the deeper cerebellar hemispheres (Lesage et al., 2012; Hardwick, Lesage, & Miall, 2014), as compared to flat figure-of-eight-shaped coils (Argyropoulos & Muggleton, 2013). On the other hand, the focality of the double-cone coil is limited as compared to that of a small figure-of-eight-shaped coil. Nevertheless, the stimulation of adjacent, motor-projecting lobules need not be an insurmountable obstacle, provided that the different experimental conditions are well-matched with respect to properties to which motor-projecting lobules may be sensitive. While lobule VIIa also includes a vermal component, known as the ‘neocerebellar vermis’, the limited size of this portion and its adjacency to paravermal regions makes it difficult to stimulate in a selective fashion. This becomes an issue, when considering the involvement of paravermal regions in articulatory processes (Petersen et al., 1989; Frings et al., 2006). Irrespective of the coil shape and size, neuronavigated localization of the stimulation site (e.g. Halko et al., 2014) would contribute significantly towards identifying the particular lobular regions of (H)VIIa that receive the greatest stimulation.

7. *Outstanding questions*

Aside from overcoming the aforementioned weaknesses in methods and the interpretations of the findings yielded so far, this newly emerging field of studies on cerebellar contributions to ‘non-motor’ aspects of predictive language processing would benefit substantially from addressing a series of outstanding questions. Some of those are outlined below:

1. *What is the connectivity?* In control-theoretic accounts of cerebellar internal models, the anatomical evidence for the connectivity of specific cerebellar lobules with specific cerebral cortical areas is available for both motor control

as well as higher cognition (Kelly & Strick, 2003). Discussions of cerebellar contributions to non-motor aspects of language processing have piggy-backed on the latter. Many authors have speculated that input to the cerebellum from Broca's area would provide an efference copy based on which the cerebellum would generate and transmit its predictions back to the prefrontal cortex (Ito, 2000a; Argyropoulos, 2009; Lesage et al., 2012). Functional connectivity studies (e.g. Buckner et al., 2011) are consistent with those ideas. However, apart from evidence on the cortico-pontine projection from area 45B (Schmahmann & Pandya, 1997) and the fact that such projections have massively expanded in humans (Ramnani et al., 2006), anatomical evidence for cerebello-dentate projections back to area 45B or for temporo-parietal cortical projections to the inferior olive remains poor (see Ito (2008) for some references).

2. *What about mirror neurons?* What is the relationship between cerebellar internal models and alleged cerebral cortical internal models that mirror neuron circuits implement (e.g. Iacoboni, 2005; Hurley, 2008)? Do cerebellar internal models output predictions by covertly employing the production system in perceiving conspecifics (e.g. Blakemore & Decety, 2001), like the alleged internal models in mirror-neurons? If so, what are the consequences for interpreting cerebellar involvement in language comprehension? Little attention has been paid to these questions so far (Miall, 2003; Strick, Dum, & Fiez, 2009).
3. *Is there a level-specific topography in the linguistic predictions of the cerebellum?* If the cerebellum is involved in predictive language processing after all, do different lobules or lobular regions generate predictions for different levels (phonology, semantics)? No study has so far identified the particular level at which these predictions occur. The effect of stimulating the posterior vermis/paravermis on phrasal predictions (Argyropoulos, 2011a) could in principle be contrasted to the effects of lateral cerebellar stimulation on semantic predictions (Lesage, 2012; Argyropoulos & Muggleton, 2013), but there are major inherent weaknesses in those studies to establish this.
4. *What about production?* The studies conducted so far have all addressed the involvement of cerebellar forward models in ‘non-motor’ aspects of prediction in language comprehension. This has been mainly driven by considerations of

confounds with articulatory motor control that studying language generation would involve. However, prediction may involve both forward and inverse internal models at different levels of both language comprehension and generation (Pickering & Garrod, 2013). Indeed, the cerebellum may encode both forward and inverse internal models (Wolpert & Kawato, 1998). Can we then study cerebellar predictions in 'non-motor' aspects of language production?

8. *Conclusion*

Recent findings on the contributions of neocerebellar circuitry to the 'non-motor' aspects of predictive language processing are highly consistent with the involvement of the cerebellum in 'higher cognitive' aspects of behavior as well as with the multi-modal nature of this unitary cerebellar computation. However, a series of significant limitations in the few studies conducted so far raise questions about the interpretation of their findings. Addressing those weaknesses will provide the opportunity to investigate whether cerebellar internal models play a role in 'non-motor' aspects of predictive language processing. Above all, such research will improve our understanding of the 'modulatory' role of the cerebellum in language processing from a computationally grounded perspective that integrates motor, cognitive and affective aspects of cerebellar function (Mariën et al., 2014).

References

Ackermann, H., Wildgruber, D., & Grodd, W. (1998). Does the cerebellum contribute to cognitive aspects of speech production? A functional MRI study in humans. *Neuroscience Letters*, *247*, 187-190.

Ackermann, H., Mathiak, K., & Ivry, R. B. (2004). Temporal organization of —internal speech as a basis for cerebellar modulation of cognitive functions. *Behavioral and Cognitive Neuroscience Reviews*, *3*, 14-22.

Ackermann, H., Mathiak, K., & Riecker, A. (2007). The contribution of the cerebellum to speech production and speech perception: Clinical and functional imaging data. *Cerebellum*, *6*, 202–213.

Albus, J.S. (1971). A theory of cerebellar function. *Mathematical Biosciences*, *10*, 25– 61.

Altmann, G.T.M., & Kamide, Y. (1999). Incremental interpretation at verbs. *Cognition*, *73*, 247–264.

Arasanz, C.P., Staines, W.R., Roy, E.A., & Schweizer, T.A. (2012). The cerebellum and its role in word generation: a cTBS study. *Cortex*, *48*(6), 718-724.

Argyropoulos, G. P. (2008). The subcortical foundations of grammaticalization. In A. D. M. Smith, K. Smith, & R. Ferrer i Cancho (Eds.). *The evolution of language: Proceedings of the 7th international conference on the evolution of language* (pp. 10–17). Singapore: World Scientific.

Argyropoulos, G. P. (2009). Neocerebellar emulation in language processing. In K. Alter, M. Horne, M. Lindgren, M. Roll, & J. von Koss Torkildsen (Eds.). *Brain Talk: Discourse with and in the brain. Papers from the first Birgit Rausing language program conference in linguistics* (pp. 193–206). Lund: Lund University, Media Tryck.

Argyropoulos, G. P. (2011a). Cerebellar theta-burst stimulation selectively enhances lexical associative priming. *Cerebellum*, *10*(3), 540-50.

Argyropoulos, G.P. (2011b). *The neocerebellar Kalman filter linguistic processor: from grammaticalization to transcranial magnetic stimulation.* Unpublished doctoral dissertation. University of Edinburgh.

https://www.era.lib.ed.ac.uk/bitstream/1842/5694/2/Argyropoulos%202011.pdf

Argyropoulos, G. P., Kimiskidis, V., & Papagiannopoulos, S. (2011). Theta-burst stimulation of the right neocerebellar vermis selectively disrupts the practice-

induced acceleration of lexical decisions. *Behavioral Neuroscience*, *125*(5), 724-34.

Argyropoulos, G.P., & Muggleton, N. (2013). Effects of cerebellar stimulation on processing semantic associations. *Cerebellum*, *12*(1), 83-96.

Balsters, J.H., Cussans, E., Diedrichsen, J., Phillips, K.A., Preuss, T.M., Rilling, J.K., & Ramnani, N. (2010). Evolution of the cerebellar cortex: the selective expansion of prefrontal-projecting cerebellar lobules. *Neuroimage*, *49*, 2045-2052

Balsters, J. H., & Ramnani, N. (2011). Cerebellar Plasticity and Automation of Rule-Related Processing *Journal of Neuroscience*, *31*(6), 2305–2312.

Balsters, J. H., & Ramnani, N. (2008). Symbolic representations of action in the human cerebellum. *Neuroimage 43*(2), 388-398.

Balsters, J.H., Whelan, C.D., Robertson, I.H., & Ramnani, N. (2013). Cerebellum and cognition: evidence for the encoding of higher order rules. *Cerebral Cortex*, *23*(6), 433-43.

Bellebaum, C., & Daum, I. (2011). Mechanisms of cerebellar involvement in associative learning. *Cortex*, *47*(1), 128–36.

Bernard, J.A., Seidler, R.D., Hassevoort, K.M., Benson, B.L., Welsh, R.C., Wiggins, J.L., Jaeggi, S.M., Buschkuehl, M., Monk, C.S., Jonides, J., & Peltier, S.J. (2012). Resting state cortico-cerebellar functional connectivity networks: a comparison of anatomical and self-organizing map approaches. *Frontiers in Neuroanatomy*, *6*(31).

Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S., Rao, S. M., & Cox, R. W. (1999). Conceptual processing during the conscious resting state. A functional MRI study. *Journal of Cognitive Neuroscience*,*11*, 80–95.

Blakemore, S.-J., & Decety, J. (2001). From the perception of action to the understanding of intention. *Neuroscience*, *2*, 561-567.

Blakemore, S.-J., Frith, C. D., & Wolpert, D. M. (2001). The cerebellum is involved in predicting the sensory consequences of action. *NeuroReport*, *12*, 1879-84.

Bloedel, J.R. (1992). Functional heterogeneity with structural homogeneity - How does the cerebellum operate? *Behavioral and Brain Sciences*, *15*, 666–678.

Blumstein, S. E., Milberg, W. P., & Shrier, R. (1982). Semantic processing in aphasia: Evidence from an auditory lexical decision task. *Brain and Language*, *17*, 301–315.

Brown, C., & Hagoort, P. (1993). The processing nature of the N400: Evidence from

masked priming. *Journal of Cognitive Neuroscience*, *5*, 34 –44.

Brysbaert, M., Warriner, A.B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, *46*, 904-911.

Buccino, G., Riggio, L., Melli, G., Binkofski, F., Gallese, V., & Rizzolatti, G. (2005). Listening to action-related sentences modulates the activity of the motor system: a combined TMS and behavioral study. *Brain Research Cognitive Brain Research*, *24*, 355–63.

Buckner, R.L., Krienen, F.M., Castellanos, A., Diaz, J.C., & Yeo, B.T.T. (2011). The organization of the human cerebellum estimated by intrinsic functional connectivity *Journal of Neurophysiology, 106*, 2322–2345.

Callan, D.E., Kawato, M., Parsons, L., & Turner, R. (2007). Speech and song: the role of the cerebellum. *Cerebellum*, *6*, 321-327.

Canavan, A. G., Sprengelmeyer, R., Diener, H. C., & Hömberg, V. (1994). Conditional associative learning is impaired in cerebellar disease in humans. *Behavioral Neuroscience*, *108*, 475-85.

Cappa, S.F., Sandrini, M., Rossini, P.M., Sosta, K., & Miniussi, C. (2002). The role of the left frontal lobe in action naming: rTMS evidence. *Neurology*, *59*, 720–3.

Chomsky, N. (1981). *Lectures on government and binding.* Dordrecht: Foris.

Christian, K. M., & Thompson, R. F. (2005). Long-term storage of an associative memory trace in the cerebellum. *Behavioral Neuroscience*, *119*, 526-37.

Coffman, K., Dum, R. P., & Strick, P. L. (2011). Cerebellar vermis is a target of projections from the motor areas in the cerebral cortex. *Proceedings of the National Academy of Sciences*, *108*(38), 16068–73.

Coltheart, M. (1981). The MRC Psycholinguistic Database. *Quarterly Journal of Experimental Psychology*, *33*A, 497-505.

Courchesne, E., & Allen, G. (1997). Prediction and preparation, fundamental functions of the cerebellum. *Learning & Memory*, *4*(1), 1-35.

Daskalakis, Z.J., Paradiso, G.O., Christensen, B.K., Fitzgerald, P.B., Gunraj, C., & Chen, R. (2004). Exploring the connectivity between the cerebellum and motor cortex in humans. *Journal of Physiology*, *557*, 689–700.

Daum, I., Ackermann, H., Schugens, M.M., Reimold, C., Dichgans, J., & Birbaumer, N. (1993). The cerebellum and cognitive functions in humans. *Behavioral Neuroscience*, *107*(3), 411–19.

De Smet, H.J., Paquier, P., Verhoeven, J., & Mariën, P. (2013). The cerebellum: its role in language and related cognitive and affective functions. *Brain & Language*, *127*, 334-342.

Desmond, J. E., & Fiez, J. A. (1998). Neuroimaging studies of the cerebellum: Language, learning, and memory. *Trends in Cognitive Sciences*, *2*(9), 355-62.

Desmond, J. E., Gabrieli, J. D. E., & Glover, G. H. (1998). Dissociation of frontal and cerebellar activity in a cognitive task: evidence for a distinction between selection and research. *Neuroimage*, *7*, 368-76.

Desmurget, M., Gréa, H., Grethe, J. S., Prablanc, C., Alexander, G. E., & Grafton, S. T. (2001). Functional anatomy of nonvisual feedback loops during reaching: a positron emission tomography study. *Journal of Neuroscience*, *21*, 2919-2928.

Devlin, J. T., & Watkins, K. E. (2007). Stimulating language: insights from TMS. *Brain*, *130*(3), 610-622.

Ding, H., Qin, W., Jiang, T., Zhang, Y., & Yu, C. (2012). Volumetric variation in subregions of the cerebellum correlates with working memory performance. *Neuroscience Letters*, *508*(1), 47–51.

Dow, R. S., & Moruzzi, G. (1958). *The physiology and pathology of the cerebellum*. Minneapolis: University of Minnesota Press.

Drepper, J., Timmann, D., Kolb, F. P., & Diener, H. C. (1999). Non-motor associative learning in patients with isolated degenerative cerebellar disease. *Brain*, *122*, 87-97.

E, K.-H., Chen, S.-H. A., Ho, M.-H. R., & Desmond, J. E. (2014). A meta-analysis of cerebellar contributions to higher cognition from PET and fMRI studies. *Human Brain Mapping*, *35*(2), 593-615.

Esposito, F., Bertolino, A., Scarabino, T., Latorre, V., Blasi, G., Popolizio, T., Tedeschi, G., Cirillo, S., Goebel, R., & Di Salle, F. (2006). Independent component model of the default-mode brain function: Assessing the impact of active thinking. *Brain Research Bulletin, 70*, 263–269.

Fedorenko, E., Hsieh, P.-J., Nieto-Castañón, A., Whitfield-Gabrieli, S., & Kanwisher, N. (2010). New method for fMRI investigations of language: defining ROIs functionally in individual subjects. *Journal of Neurophysiology*, *104*(2), 1177–94.

Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology*, *47*, 164-203.

Ferrucci, R., Marceglia, S., Vergari, M., Cogiamanian, F., Mrakic-Sposta, S., Mameli,

F., Zago, S., Barbieri, S., & Priori, A. (2008). Cerebellar transcranial direct current stimulation impairs the practice-dependent proficiency increase in working memory. *Journal of Cognitive Neuroscience*, *20*(9), 1687-97.

Fiez, J. A., & Raichle, M. (1997). Linguistic processing. *International Review of Neurobiology, 41*, 233-54.

Fiez, J.A., Petersen, S.E., Cheney, M.K., & Raichle, M.E. (1992). Impaired nonmotor learning and error detection associated with cerebellar damage. A single case study. *Brain*, *115*, 155–78.

Fransson, P. (2005). Spontaneous low-frequency BOLD signal fluctuations: an fMRI investigation of the resting-state default mode of brain function hypothesis. *Human Brain Mapping, 26*, 15–29.

Fransson, P. (2006). How default is the default mode of brain function? further evidence from intrinsic BOLD signal fluctuations. *Neuropsychologia*, *44*, 2836–2845.

Friederici, A. D. (2012). The cortical language circuit: from auditory perception to sentence comprehension. *Trends in Cognitive Sciences*, *16*(5), 262–8.

Frings, M., Dimitrova, A., Schorn, C.F., Elles, H.-G., Hein-Kropp, C., Gizewski, E.R., Diener, H.C., & Timmann, D. (2006). Cerebellar involvement in verb generation: an fMRI study. *Neuroscience Letters*, *409*, 19–23.

Gebhart, A. L., Petersen, S. E., & Thach, W. T. (2002). Role of the posterolateral cerebellum in language. *Annals of the New York Academy of Science*, *978*, 318-33.

Ghosh S.S., Tourville, J.A., & Guenther F.H. (2009). A neuroimaging study of premotor lateralization and cerebellar involvement in the production of phonemes and syllables. *Journal of Speech, Language, and Hearing Research*, *51*(5), 1183-1202.

Gilbert, P.F. & Thach, W.T. (1977). Purkinje cell activity during motor learning. *Brain Research*, *128*, 309–328.

Glaser, Y.G., Martin, R.C., Van Dyke, J.A., Hamilton, A.C., & Tan, Y. (2013). Neural basis of semantic and syntactic interference in sentence comprehension. *Brain & Language*, *126*, 314–326.

Golfinopoulos, E., Tourville, J. a, Bohland, J. W., Ghosh, S. S., Nieto-Castanon, A., & Guenther, F. H. (2011). fMRI investigation of unexpected somatosensory feedback perturbation during speech. *Neuroimage*, *55*(3), 1324–38.

Golfinopoulos, E., Tourville, J., & Guenther, F. H. (2010). The integration of large-scale neural network modeling and functional brain imaging in speech motor control. *Neuroimage*, *52*(3), 862–74.

Greicius, M.D., & Menon, V. (2004). Default-mode activity during a passive sensor task: Uncoupled from deactivation but impacting activation. *Journal of Cognitive Neuroscience*, *16*, 1484–1492.

Grimaldi, G., Argyropoulos, G.P., Bastian, A., Cortes, M., Davis, N.J., Edwards, D., Ferrucci, R., Fregni, F., Galea, J.M., Hamada, M., Manto, M., Miall, R.C., Morales-Quezada, L., Pope, P.A., Priori, A., Rothwell, J., Tomlinson, S.P., & Celnik, P. (2014). Cerebellar transcranial direct current stimulation (ctDCS): a novel approach to understand cerebellar function in health and disease. *Neuroscientist* [epub ahead of print].

Grimaldi, G., Argyropoulos, G.P., Boehringer, A., Celnik, P., Edwards, M.J., Ferrucci, R., Galea, J.M., Groiss, S.J., Hiraoka, K., Kassavetis, P., Lesage, E., Manto, M., Miall, R.C., Priori, A., Sadnicka, A., Ugawa, Y., & Ziemann, U. (2013). Non-invasive cerebellar stimulation—a consensus paper. *Cerebellum*, *13*(1), 121-138.

Grush, R. (2004). The emulation theory of representation: motor control, imagery, and perception. *Behavioral and Brain Sciences*, *27*, 377-435.

Guediche, S., Holt, L. L., Laurent, P., Lim, S.-J., & Fiez, J. A. (2014). Evidence for cerebellar contributions to adaptive plasticity in speech perception. *Cerebral Cortex*. [Epub ahead of print].

Gusnard, D.A., & Raichle, M.E. (2001). Searching for a baseline: functional imaging and the resting human brain. *Nature Reviews Neuroscience, 2*, 685–694.

Gusnard, D.A., Akbudak, E., Shulman, G.L., & Raichle, M.E. (2001). Medial prefrontal cortex and self-referential mental activity: relation to a default mode of brain function. *Proceedings of the National Academic of Sciences*, 98(7), 4259–4264.

Habas, C., Kamdar, N., Nguyen, D., Prater, K., Beckmann, C. F., Menon, V., & Greicius, M. D. (2009). Distinct cerebellar contributions to intrinsic connectivity networks. *Journal of Neuroscience*, *29*(26), 8586–94.

Hagoort, P., Baggio, G., & Willems, R. M. (2009). Semantic unification. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences*, 4th ed. (pp. 819-836). Cambridge, MA: MIT Press.

Halko, M.A., Farzan, F., Eldaief, M.C., Schmahmann, J.D., & Pascual-Leone, A.

(2014). Intermittent theta-burst stimulation of the lateral cerebellum increases functional connectivity of the default network. *Journal of Neuroscience, 34*(36), 12049-12056.

Hanakawa, T., Dimyan, M.A., & Hallett, M. (2008). Motor planning, imagery, and execution in a distributed motor network: a time-course study with functional MRI. *Cerebral Cortex*, *18*, 2775–88.

Hardwick, R.M., Lesage, E., & Miall, R.C. (2014). Cerebellar transcranial magnetic stimulation: the role of coil geometry and tissue depth. *Brain Stimulation*, *7*, 643-649.

Harrison, B.J., Pujol, J., López-Solà, M., Hernández-Ribas, R., Deus, J., Ortiz, H., Soriano-Mas, C., Yücel, M., Pantelis, C., & Cardoner, N. (2008). Consistency and functional specialization in the default mode brain network. *Proceedings of the National Academic of Sciences*, *105*(28), 9781–9786.

Hashimoto, M., & Ohtsuka, K. (1995). Transcranial magnetic stimulation over the posterior cerebellum during visually guided saccades in man. *Brain*, *118*, 1185–1193.

Hauk, O., Johnsrude, I., & Pulvermuller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron*, *41*, 301–7.

Heinks-Maldonado, T. H., Nagarajan, S. S., & Houde, J. F. (2006). Magnetoencephalographic evidence for a precise forward model in speech production. *Neuroreport*, *17*(13), 1375–1379.

Hickok, G. (2012). Computational neuroanatomy of speech production. *Nature Reviews Neuroscience*, *13*, 135–145.

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, *8*, 393–402.

Hosemann, J., Herrmann, A., Steinbach, M., Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2013). Lexical prediction via forward models: N400 evidence from German Sign Language *Neuropsychologia*, *51*, 2224–2237.

Hurley, S. (2008). The shared circuits model (SCM): How control, mirroring, and simulation can enable imitation, deliberation, and mindreading. *Behavioral and Brain Sciences*, 31, 1- 58.

Hutchison, K. A. (2003). Is semantic priming due to association strength or feature overlap? A *micro*analytic review. *Psychonomic Bulletin & Review*, *10*(4), 785-813.

Iacoboni, M. (2008). The role of premotor cortex in speech perception: evidence from

fMRI and rTMS. *Journal of Physiology*, *102*(1-3), 31-34.

Imamizu, H, Miyauchi, S., Tamada, T., Sasaki, Y., Takino, R., Pütz, B., Yoshioka, T., & Kawato, M. (2000). Human cerebellar activity reflecting an acquired internal model of a new tool. *Nature*, *403*, 192-5.

Ito, M. (1970). Neurophysiological basis of the cerebellar motor control system. *International Journal of Neurology*, *7*, 162-76.

Ito, M. (1984). *The cerebellum and neural control.* New York: Raven Press.

Ito, M. (2000a). Neural control of cognition and language. In A. Marantz, Y. Miyashita, & W. O'Neil (Eds.), *Image, language, brain* (pp. 149-62). Cambridge, Massachusetts: MIT Press.

Ito, M. (2000b). Neurobiology: internal model visualized. *Nature*, *403*, 153-154.

Ito, M. (2008). Control of mental activities by internal models in the cerebellum. *Nature Reviews Neuroscience*, *9*, 304-13.

Jones, L.L., & Estes, Z. (2012). Lexical Priming: associative, semantic, and thematic influences on word recognition. In: J. Adelman (Ed.), *Visual Word Recognition, Vol II: Meaning and Context, Individuals and Development* (pp. 44-72). Psychology Press.

Jordan, M. I., & Wolpert, D. M. (2000). Computational motor control. In M. S. Gazzaniga, (Ed.), *The new cognitive neurosciences* (pp. 601-618). Cambridge, MA: MIT Press.

Jueptner, M., Frith, C.D., Brooks, D.J., Frackowiak, R.S.J., & Passingham, R.E. (1997). Anatomy of motor learning. II. Subcortical structures and learning by trial and error. *Journal of Neurophysiology*, *77*, 1325-37.

Kaan, E., & Swaab, T.Y. (2002). The brain circuitry of syntactic comprehension. *Trends in Neurosciences*, *6*(8), 350-356.

Kawato, M. (1999). Internal models for motor control and trajectory planning, *Current Opinion in Neurobiology*, *9*, 718-27.

Kawato, M., Kuroda, T., Imamizu, H., Nakano, E., Miyauchi, S., & Yoshioka, T. (2003). Internal forward models in the cerebellum: fMRI study on grip force and load force coupling. *Progress in Brain Research*, *142*, 171-88.

Kelly, R.M., & Strick, P.L. (2003). Cerebellar loops with motor cortex and prefrontal cortex of a nonhuman primate. *Journal of Neuroscience*, *23(23)*, 8432–44.

Kincses, Z.T., Johansen-Berg, H., Tomassini, V., Bosnell, R., Matthews, P.M., & Beckmann, C.F. (2008). Model-free characterization of brain functional networks

for motor sequence learning using fMRI. *Neuroimage*, *39*(4), 1950–1958.

Kotz, S.A., & Schwartze, M. (2010). Cortical speech processing unplugged: a timely subcortico-cortical framework. *Trends in Cognitive Sciences*, *14*, 392-399.

Krienen, F.M., & Buckner, R.L. (2009). Segregated fronto-cerebellar circuits revealed by intrinsic functional connectivity. *Cerebral cortex*, *19*, 2485-2497.

Lang, C. E., & Bastian, A. J. (2002). Cerebellar damage impairs automaticity of a recently practiced movement. *Journal of Neurophysiology*, *87*(3), 1336-47.

Leggio, M.G., Silveri, M.C., Petrosini, L., & Molinari, M. (2000). Phonological grouping is specifically affected in cerebellar patients: a verbal fluency study. *Journal of Neurology, Neurosurgery, and Psychiatry*, *69*, 102–6.

Leggio, M.G., Tedesco, A. M., Chiricozzi, F. R., Clausi, S., Orsini, A., & Molinari, M. (2008). Cognitive sequencing impairment in patients with focal or atrophic cerebellar damage. *Brain*, *131*, 1332-1343.

Leiner, H.C. (2010). Solving the mystery of the human cerebellum. *Neuropsychology Review*, *20*, 229-235.

Lesage, E. (2013). *On the cerebellum and language: neurostimulation and imaging studies.* Unpublished doctoral dissertation. University of Birmingham. http://etheses.bham.ac.uk/4926/1/Lesage14PhD.pdf

Lesage, E., Hansen, P.C., & Miall, C.C. (2014). Cerebellar BOLD response to linguistic stimuli is modulated by predictability. *Society for Neuroscience abstracts.*

Lesage, E., Morgan, B.E., Olson, A.C., Meyer, A.S., & Miall, R.C. (2012). Cerebellar rTMS disrupts predictive language processing, *Current Biology*, *22*(18), 794-795.

Lieberman, P. (1963). Some effects of semantic and grammatical context on the production and perception of speech. *Language and Speech*, *6*, 172-187.

Liuzzi, G., Freundlieb, N., Ridder, V., Hoppe, J., Heise, K, Zimerman, M., Dobel, C., Enriquez-Geppert, S., Gerloff, C., Zwitserlood, P., & Hummel, F.C. (2010). The involvement of the left motor cortex in learning of a novel action word lexicon. *Current Biology*, *20*(19), 1745-51.

Maddox, W.T., Aparicio, P., Marchant, N.L., & Ivry, R.B. (2005). Rule-based category learning is impaired in patients with Parkinson's disease but not in patients with cerebellar disorders. *Journal of Cognitive Neuroscience*, *17*(5), 707-23.

Manto, M., Bower, J.M., Conforto, A.B., Delgado-García, J.M., da Guarda, S.N., Gerwig, M., Habas, C., Hagura, N., Ivry, R.B., Mariën, P., Molinari, M., Naito, E., Nowak, D.A., Oulad Ben Taib, N., Pelisson, D., Tesche, C.D., Tilikete, C., &

Timmann, D. (2012). Consensus paper: roles of the cerebellum in motor control--the diversity of ideas on cerebellar involvement in movement. *Cerebellum*, *11*(2), 457-87.

Mariën, P., Ackermann, H., Adamaszek, M., Barwood, C.H.S., Beaton, A., Desmond, J., De Witte, E., Fawcett, A.J., Hertrich, I., Küper, M., Leggio, M., Marvel, C., Molinari, M., Murdoch, B.E., Nicolson, R.I., Schmahmann, J.D., Stoodley, C.J., Thürling, M., Timmann, D., Wouters, E., & Ziegler, W. (2014). Consensus paper: language and the cerebellum: an ongoing enigma. *Cerebellum*, *13*, 386–410.

Mariën, P., Engelborghs, S., Fabbro, F., & De Deyn, P. P. (2001). The lateralized linguistic cerebellum: a review and a new hypothesis. *Brain and Language*, *79*, 580-600.

Marr, D. (1969). A theory of cerebellar cortex. *Journal of Physiology*, *202*, 437– 470.

Mazoyer, B., Zago, L., Mellet, E., Bricogne, S., Etard, O., Houde, O., Crivello, F., Joliot, M., Petit, L., & Tzourio-Mazoyer, N. (2001). Cortical networks for working memory and executive functions sustain the conscious resting state in man. *Brain Research Bulletin*, *54*, 287–298.

McCormick, D. A, & Thompson, R. F. (1984). Cerebellum: essential involvement in the classically conditioned eyelid response. *Science*, *223*, 296-9.

McKiernan, K.A., Kaufman, J.N., Kucera-Thompson, J., & Binder, J.R. (2003). A parametric manipulation of factors affecting task-induced deactivation in functional neuroimaging. *Journal of Cognitive Neuroscience, 15*, 394–408.

McRae, K., Hare, M., Elman, J. L., & Ferretti, T. (2005). A basis for generating expectancies for verbs from nouns. *Memory and Cognition*, *33*(7), 1174-84.

Medina, J.F. & Lisberger, S.G. (2009). Encoding and decoding of learned smooth-pursuit eye movements in the floccular complex of the monkey cerebellum, *Journal of Neurophysiology*, *102*, 2039-2054.

Miall, R. C. (2003). Connecting mirror neurons and forward models. *NeuroReport*, *14*(16), 1-3.

Miall, R. C., & Christensen, L. O. D. (2004). The effect of rTMS over the cerebellum in normal human volunteers on peg-board movement performance. *Neuroscience Letters*, 371, 185-189.

Miall, R. C., Christensen, L. O. D., Cain, O., & Stanley, J. (2007). Disruption of state estimation in the human lateral cerebellum. *Public Library of Science Biology*, *5*(11), 2733-44.

Miall, R. C., & King, D. (2008). State estimation in the cerebellum. *Cerebellum*, *7*, 572-6.

Middleton, F.A. & Strick, P.L. (2000). Basal ganglia and cerebellar loops: motor and cognitive circuits. *Brain Research Reviews*, *31*, 236-250.

Middleton, F. A., & Strick, P. L. (2001). Cerebellar projections to the prefrontal cortex of the primate. *Journal of Neuroscience*, *21*(2), 700-12.

Milberg, W. P., & Blumstein, S. E. (1981). Lexical decision and aphasia: evidence for semantic processing. *Brain and Language*, *14*, 371–385.

Moberget, T., Gullesen, E.H., Andersson, S., Ivry, R.B., & Endestad, T. (2014). Generalized role for the cerebellum in encoding internal models: evidence from semantic processing. *Journal of Neuroscience*, *34*(8), 2871–2878.

Müller, F., & Dichgans, J. (1994). Dyscoordination of pinch and lift forces during grasp in patients with cerebellar lesions. *Experimental Brain Research*, *101*, 485-92.

Nestor, P., Valdman, O., Niznikiewicz, M., Spencer, K., McCarley, R., & Shenton, M. (2006). Word priming in schizophrenia: associational and semantic influences *Schizophrenia Research*, *82*(2), 139-42.

Nowak, D. A., Timmann, D., & Hermsdorfer, J. (2007). Dexterity in cerebellar agenesis. *Neuropsychology*, *45*, 696-703.

Oliveri, M., Koch, G., Torriero, S., & Caltagirone, C. (2005) Increased facilitation of the primary motor cortex following 1 Hz repetitive transcranial magnetic stimulation of the contralateral cerebellum in normal humans. *Neuroscience Letters*, *376*, 188–193.

Oliveri, M., Bonnì, S., Turriziani, P., Koch, G., Gerfo, E. L., Torriero, S., Vicario, C.M., Petrosini, L. & Caltagirone, C. (2009). Motor and linguistic linking of space and time in the cerebellum. *PloS One*, *4*(11), e7933.

Perkell, J. S., Matthies, M. L., Lane, H., Guenther, F. H., Wilhelms-Tricarico, R., Wozniak, J. & Guiod, P. (1997). Speech motor control: acoustic goals, saturation effects, auditory feedback and internal models. *Speech Communication*, *22*, 227-250.

Persson, J., Lustig, C., Nelson, J., & Reuter-Lorenz, P. (2007). Age differences in deactivation: a link to cognitive control? *Journal of Cognitive Neuroscience*, *19*, 1021–1032.

Petersen, S. E., Fox, P. T., Posner, M. L., Mintun, M., & Raichle, M. E. (1989).

Positron emission tomographic studies of the processing of single words. *Journal of Cognitive Neuroscience*, *1*, 153-70.

Picard, H., Amado, I., Mouchet-Mages, S., Olié, J.- P., & Krebs, M.-O. (2008). The role of the cerebellum in schizophrenia: an update of clinical, cognitive, and functional evidences. *Schizophrenia Bulletin*, *34*(1), 155-172.

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, *27*, 169-225.

Pickering, M. J., & Garrod, S. (2007). Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences*, *11*(3), 105-10.

Pickering, M. J., & Garrod, S. (2013). Forward models and their implications for production, comprehension, and dialogue. *Behavioral and Brain Sciences*, *36*(4), 377-392.

Poeppel, D., Emmorey, K., Hickok, G., & Pylkkänen, L. (2012). Towards a new neurobiology of language. *Journal of Neuroscience*, *32*,14125–14131.

Polli, F., Barton, J., Cain, M., Thakkar, K., Rauch, S., & Manoach, D. (2005). Rostral and dorsal anterior cingulate cortex make dissociable contributions during antisaccade error commission. *Proceedings of the National Academy of Science*, *102*, 15700–15705.

Pope, P., & Miall, R.C. (2012). Task-specific facilitation of cognition by cathodal transcranial direct current stimulation of the cerebellum. *Brain Stimulation*, *5*, 84-94.

Pulvermüller, F., Hauk, O., Nikulin, V.V., & Ilmoniemi, R.J. (2005). Functional links between motor and language systems. *European Journal of Neuroscience*, *21*, 793–7.

Puttemans, V., Wenderoth, N., & Swinnen, S.P. (2005). Changes in brain activation during the acquisition of a multifrequency bimanual coordination task: from the cognitive stage to advanced levels of automaticity. *Journal of Neuroscience, 25*, 4270–4278.

Raboyeau, G., Marie, N., Balduyck, S., Gros, H., Démonet, J. -F., & Cardebat, D. (2004). Lexical learning of the English language: a PET study in healthy French subjects. *Neuroimage*, *22*(4), 1808–18.

Raichle, M.E., & Gusnard, D.A. (2005). Intrinsic brain activity sets the stage for expression of motivated behaviour. *Journal of Comparative Neurology*, *493*(1), 167-176.

Raichle, M.E., MacLeod, A.M., Snyder, A.Z., Powers, W.J., Gusnard, D.A., & Shulman, G.L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences*, *98*, 676–682.

Ramnani, N. (2006). The primate cortico-cerebellar system: anatomy and function. *Nature Reviews Neuroscience*, *7*, 511–522.

Ramnani, N., Behrens, T.E., Johansen-Berg, H., Richter, M.C., Pinsk, M.A., Andersson, J.L., Rudebeck, P., Ciccarelli, O., Richter, W., Thompson, A.J., Gross, C.G., Robson, M.D., Kastner, S., & Matthews, P.M. The evolution of prefrontal inputs to the cortico-pontine system: diffusion imaging evidence from macaque monkeys and humans (2006). *Cerebral Cortex*, *16*(6), 811-8.

Rao, N.K., Motes, M.A., & Rypma, B. (2014). Investigating the neural bases for intra-subject cognitive efficiency changes using functional resonance imaging. *Frontiers in Human Neuroscience*, *8*, 840.

Riva, D., & Giorgi, C. (2000). The cerebellum contributes to higher functions: evidence from a series of children surgically treated for posterior fossa tumours. *Brain*, *123*, 1051-1061.

Rogers, S. L., & Friedman, R. B. (2008). The underlying mechanisms of semantic memory in Alzheimer’s disease and semantic dementia. *Neuropsychologia*, *46*, 12-21.

Roxbury, T., McMahon, K., & Copland, D.A. (2014). An fMRI study of concreteness effects in spoken word recognition. *Behavioral and Brain Functions*, *10*(34),1-14.

Schmahmann, J. D. (2000). The role of the cerebellum in affect and psychosis. *Journal of Neurolinguistics*, *13*, 189-214.

Schmahmann, J. D., & Pandya, D. N. (1997). The cerebrocerebellar system. *International Review of Neurobiology*, *41*, 31-60.

Shapiro, K.A., Pascual-Leone, A., Mottaghy, F.M., Gangitano, M., & Caramazza, A. (2001). Grammatical distinctions in the left frontal cortex. *Journal of Cognitive Neuroscience*, *13*, 713–20.

Shulman, G.L., Fiez, J.A., Corbetta, M., Buckner, R.L., Miezin, F.M., Raichle, M.E., & Petersen, S.E. (1997). Common blood flow changes across visual tasks: I. Increases in subcortical structures and cortex. *Journal of Cognitive Neuroscience*, *9*, 648–663.

Singh, K.D., & Fawcett, I.P. (2008). Transient and linearly graded deactivation of the human default-mode network by a visual detection task. *Neuroimage*, *41*,100-112.

Staub, A., & Clifton, C. Jr. (2006). Syntactic prediction in language comprehension: evidence from either…or. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(2), 425-436.

Stoodley, C.J., & Schmahmann, J.D. (2009). Functional topography in the human cerebellum: a meta-analysis of neuroimaging studies. *Neuroimage, 44*(2), 489-501.

Strick P.L., Dum, R.P., & Fiez, J.A. (2009). Cerebellum and nonmotor function. *Annual Review of Neuroscience*, *32*, 413–434.

Sturt, P., & Lombardo, V. (2005). Processing coordinated structures: incrementality and connectedness. *Cognitive Science*, *29*, 291-305.

Thach, W. T. (1997). Context-response linkage. *International Review of Neurobiology*, *41*, 599-611.

Thompson, R. F., Bao, S., Chen, L., Cipriano, B. D., Grethe, J. S., Kim, J. J., Thompson, J. K., Tracy, J. A., Weninger, M. S., & Krupa, D. J. (1997). Associative learning. *International Review of Neurobiology*, *41*, 152-89.

Thompson-Schill, S. L., Kurtz, K. J., & Gabrieli, J. D. E. (1998). Effects of semantic and associative relatedness on automatic priming. *Journal of Memory and Language*, *38*, 440-58.

Tian, X., & Poeppel, D. (2010). Mental imagery of speech and movement implicates the dynamics of internal forward models. *Frontiers in Psychology*, *1*, 1–23.

Timmann, D., Drepper, J., Calabrese, S., Bürgerhoff, K., Maschke, M., Kolb, F.P., Daum, I., & Diener, H. C. (2004). Use of sequence information in associative learning in control subjects and cerebellar patients. *Cerebellum*, 3, 75–82.

Timmann, D., Drepper, J., Frings, M., Maschke, M., Richter, S., Gerwig, M., & Kolb, F.P. (2010). The human cerebellum contributes to motor, emotional and cognitive associative learning. A review. *Cortex*. *46*, 845–57.

Timmann, D., Drepper, J., Maschke, M., Kolb, F.P., Böring, D., Thilmann, A.F., & Diener, H.C. (2002). Motor deficits cannot explain impaired cognitive associative learning in cerebellar patients. *Neuropsychologia*, *40*, 788–800.

Timmann, D., Kolb, F. P., Baier, C., Rijntjes, M., Mueller, S. P., Diener, H. C., & Weiller, C. (1996). Cerebellar activation during classical conditioning of the human flexion reflex: a PET study. *NeuroReport*, *7*, 2056-60.

Toni, I., Krams, M., Turner, R., & Passingham, R. E. (1998). The time course of changes during motor sequence learning: a whole-brain fMRI study. *Neuroimage*,

8, 50-61.

Tourville, J., Reilly, K. J., & Guenther, F. H. (2008). Neural mechanisms underlying auditory feedback control of speech. *Neuroimage*, *39*(3), 1429–43.

Townsend, D., & Bever, T. G. (2001). *Sentence comprehension: the integration of habits and rules.* Cambridge, Massachusetts: MIT Press.

Tucker, J., Harding, A. E., Jahanshahi, M., Nixon, P. D., Rushworth, M., Quinn, N. P., Thompson, P. D., & Passingham, R. E. (1996). Associative learning in patients with cerebellar ataxia. *Behavioral Neuroscience*, *110*, 1229-34.

Watkins, K., & Paus, T. (2004). Modulation of motor excitability during speech perception: the role of Broca's area. *Journal of Cognitive Neuroscience*, *16*(6), 978-987.

Wilson, M., & Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psychological Bulletin*, *131*(3), 460-473

Wolpert, D.M., & Kawato, M. (1998). Multiple paired forward and inverse models for motor control. *Neural Networks*, *11*, 1317–1329.

Wolpert, D.M., Miall, R.C., & Kawato, M. (1998). Internal models in the cerebellum. *Trends in Cognitive Sciences*, *2*(9), 338–47.

Figure Legends

Figure 1. Neural implementation of internal (forward) models in the cerebellum; left: an internal model implemented in cerebellar lobules HV,HVI,HVIIb,HVIII interacting with the motor cortex; right: an internal model implemented in cerebellar lobule (H)VIIa Crus I/II interacting with the prefrontal cortex. Its organization is explicitly based on the well-studied cerebro-cerebellar circuitry and corresponding control-theoretic accounts of motor control (left). A copy of a process involving the interaction of the prefrontal cortex with another brain area is sent via the pontine nuclei to its interconnected cerebellar lobules. Predictions generated from those are transmitted from the Purkinje cells of the cerebellar cortex via the ventrolateral neodentate nucleus and the thalamus back to those prefrontal areas. Predicted and actual consequences of the process copied by these cerebellar lobules are compared in the inferior olive. Any discrepancies between the two are fed via climbing fibers to the cerebellar cortex in the form of an error signal. Long-term depression is triggered at the parallel fiber-to-Purkinje cell synapses, updating the internal model. RN: red nucleus. Modified with permission from Macmillan Publishers Ltd: NATURE REVIEWS NEUROSCIENCE (Ramnani, N. (2006). The primate cortico-cerebellar system: anatomy and function. Nat. Rev. Neurosci., 7, 511–522), copyright (2006).

Figure 2. Schematic representation of an internal (forward) model involved in the comprehension of the sentence ‘Harry went out to fly his red flag’ for three different levels of prediction (Pickering & Garrod, 2007). This ‘emulator’ (Grush, 2004) is controlled by feedback from a Kalman filter, weighing predictions against analysis of

the speech input at each (arbitrarily defined here) step. In case of strong predictions and/or noisy inputs, there is ‘low Kalman gain’ (strong top-down influence on the analysis); in case of poor predictions and clear input, there is ‘high Kalman gain’ (strong bottom-up influence). For the sake of the illustration, the probabilistic nature of the predictions and the interaction among predictions at different levels are not discussed. The implementation of the internal model in the language production system will not be discussed here. Figure reproduced from Pickering & Garrod (2007, p. 108) with permission © 2007 Elsevier

Figure 3. Schematic illustration of the phonological aspects of lexical prediction generated by a cerebellar internal model. Its organization is explicitly based on the well-studied cerebro-cerebellar circuitry and corresponding control-theoretic accounts of motor control (figure 1 above). A copy of a process involving the interaction of the prefrontal cortex with another brain area (such as the temporo-parietal cortex) in the context of comprehending a sentence is sent via the pontine nuclei to its interconnected cerebellar lobules. Cerebellar predictions are transmitted from the Purkinje cells of the cerebellar cortex via the ventrolateral neodentate nucleus and the thalamus back to cortical areas. Predicted and actual consequences of the process are compared in the inferior olive. Any discrepancies between the two are fed via climbing fibers from the inferior olive to the cerebellar cortex in the form of an error signal. LTD (long-term depression) is triggered at the parallel fiber-to-Purkinje cell synapses, updating the internal model. For the sake of the illustration, the cortico-cortical interactions are assumed to proceed in a bottom-up fashion (see section 7 for more discussion). See also Ito(2008) for discussion of available evidence from anatomical connectivity studies supporting these models. Same

simplifications are made as in Pickering and Garrod (2007; figure 2 here).

Table Legends

Table 1. Concreteness, imageability, and meaningfulness ratings from Coltheart (1981) and Brysbaert, Warriner, and Kuperman (2014) for the available verbs used in the ‘non-predictive’ and ‘predictive’ sentences in Lesage et al. (2012). Verbs in ‘predictive’ sentences are more concrete, imageable, and meaningful as compared to those in ‘non-predictive’ sentences. Stimuli retrieved from Lesage (2013).

Table 2. Peak activation coordinates (MNI) in posterolateral regions of HVIIa Crus I/II for: the contrast ‘Congruent > Scrambled’ in Moberget et al. (2014); the functional connectivity study of Krienen and Buckner (2009); contrasts in a selection of fMRI studies on default-mode network. The coordinates of Rao, Motes, and Rypma (2014) are converted from Talairach to MNI space.

Table 1

Metrics	**Non-Predictive Condition**			**Predictive Condition**			**Contrast**		
	Verbs available	**Mean rating**	**SEM rating**	**Verbs available**	**Mean rating**	**SEM rating**	***t***	***p* two-tailed**	**SE diff.**
Concreteness (Coltheart, 1981)	23/32	421.83	16.31	22/32	483.05	15.08	2.75	.009	22.27
Concreteness (Brysbaert, Warriner, & Kuperman, 2014)	32/32	3.55	.14	32/32	4.13	.08	3.66	.0006	.16
Imageability (Coltheart, 1981)	23/32	458.30	12.21	22/32	519.00	16.75	2.95	.005	20.59
Meaningfulness (Coltheart, 1981)	17/32	444.94	8.58	19/32	477.42	9.14	2.57	.015	12.62

Table 2

Study	**Description**	**Coordinates (MNI)**		
		x (mm)	**y (mm)**	**z (mm)**
Moberget et al. (2014)	‘Congruent > Scrambled’ sentences	32	-74	-37
Krienen & Buckner (2009)	Functional connectivity with medial prefrontal cortex	34	-80	-36
		-30	-78	-34
		-32	-76	-34
		24	-80	-32
Gusnard et al. (2001)	Internally > Externally Cued Judgments	25	-75	-30
Rao, Motes, & Rypma (2014)	0 > ‘Digit-Symbol Verification Task’	19	-83	-36
Harrison et al. (2008)	Rest > 0	-33	-77	-27
	Moral dilemma > 0	24	-74	-29
		-31	-71	-30
	0 > Stroop task	-33	-77	-22
Fransson (2005)	Resting-state / condition with eyes closed > 0	-34	-82	-34
		22	-84	-34
Fransson (2006)	rest > two-back working memory task	39	-72	-42

Figure 1

Click here to download high resolution image

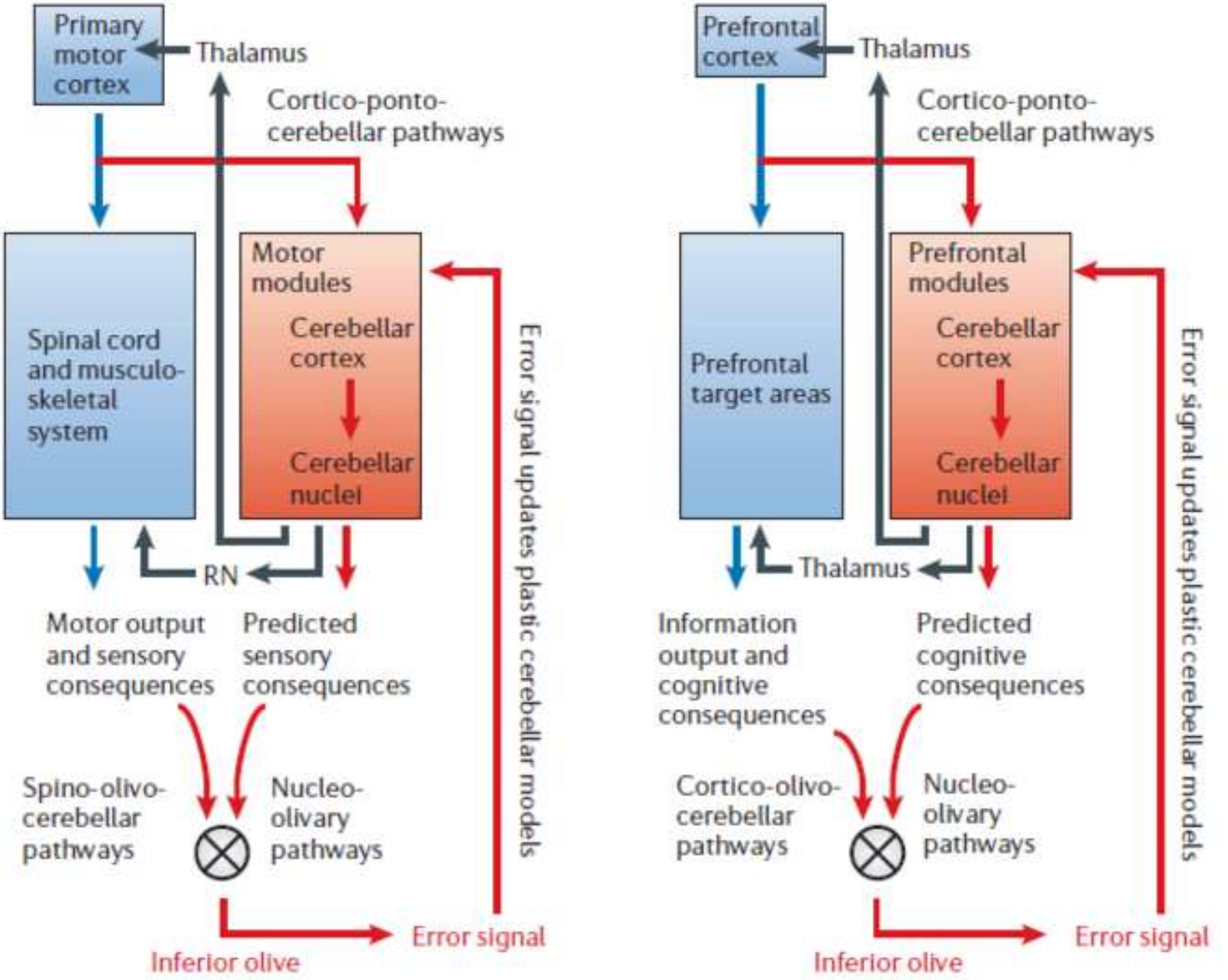
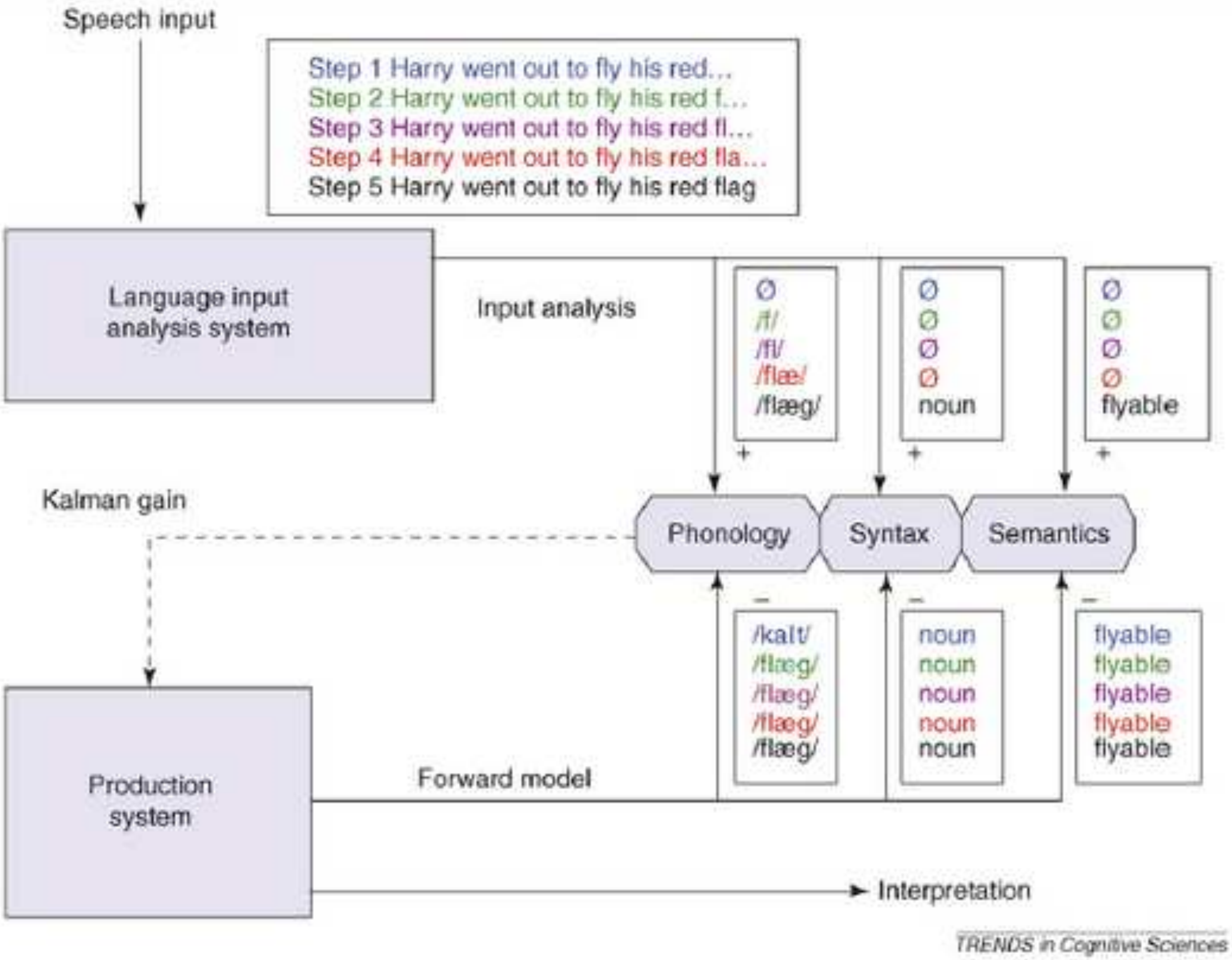


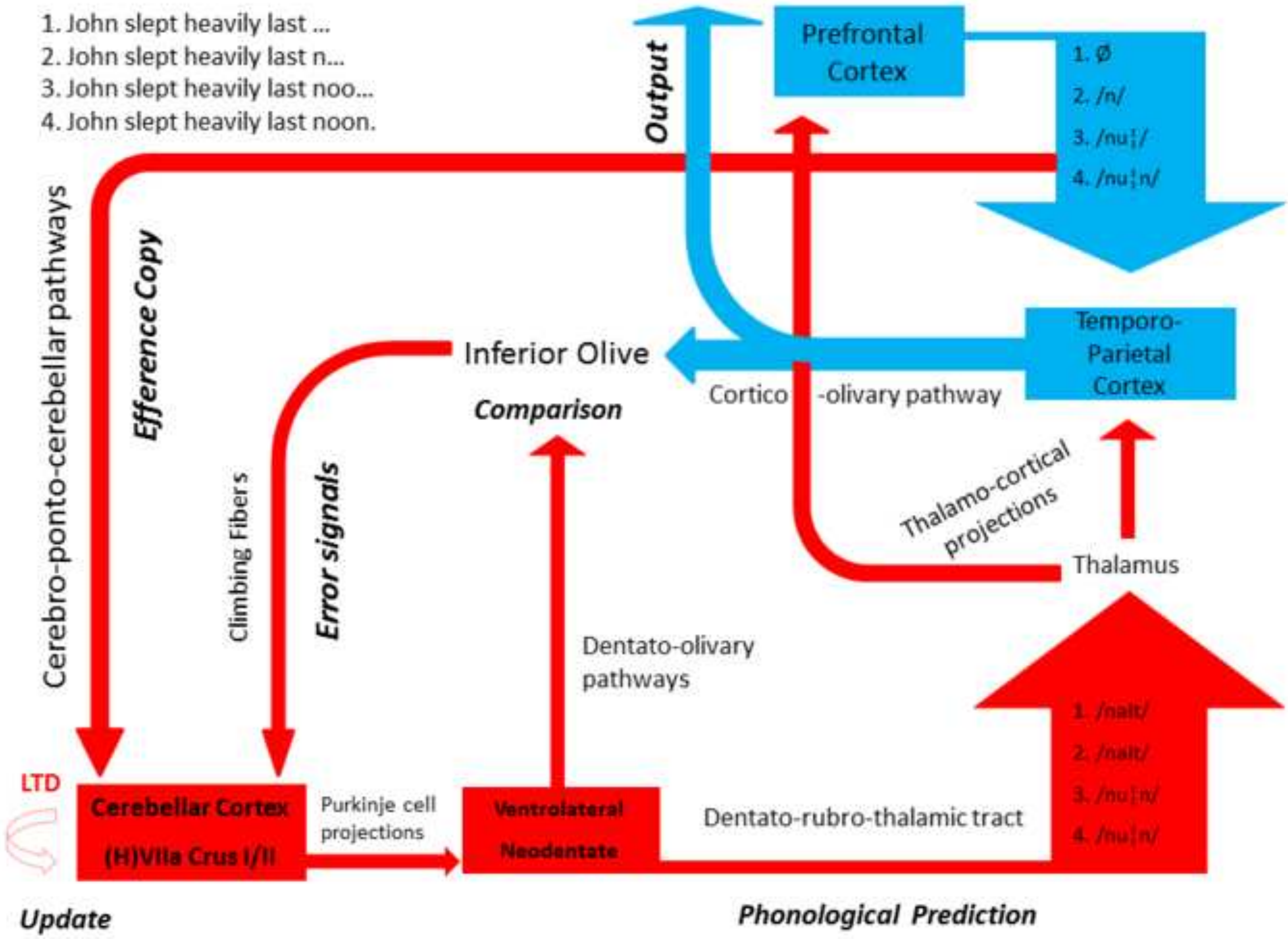
Figure 2



Step 1 Harry went out to fly his red…
Step 2 Harry went out to fly his red f…
Step 3 Harry went out to fly his red fl…
Step 4 Harry went out to fly his red fla…
Step 5 Harry went out to fly his red flag

TRENDS in Cognitive Sciences

Figure 3


1. John slept heavily last ...
2. John slept heavily last n...
3. John slept heavily last noo...
4. John slept heavily last noon.
Output
Prefrontal Cortex
1. Ø
2. /n/
3. /nu¦/
4. /nu¦n/
Cerebro-ponto-cerebellar pathways
Efference Copy
Inferior Olive
Comparison
Cortico -olivary pathway
Temporo-Parietal Cortex
Climbing Fibers
Error signals
Thalamo-cortical projections
Thalamus
Dentato-olivary pathways
1. /nait/
2. /nait/
3. /nu¦n/
4. /nu¦n/
LTD
Cerebellar Cortex
(H)VIIa Crus I/II
Purkinje cell projections
Ventrolateral
Neodentate
Dentato-rubro-thalamic tract
Update
Phonological Prediction