The Neocerebellar Kalman Filter Linguistic Processor: From Grammaticalization to Transcranial Magnetic Stimulation

Volume I

Giorgos P. Argyropoulos
B.A. (Hons.)

A thesis submitted in fulfillment of requirements for the degree of Doctor of Philosophy

to

Linguistics and English Language
School of Philosophy, Psychology and Language Sciences
University of Edinburgh
August, 2010

To Talmy Givón,

for the honor,

and to Tassos Christidis,

for the memories.

Abstract

The present work introduces a synthesis of neocerebellar state estimation and feedforward control with multi-level language processing. The approach combines insights from clinical, imaging, and modelling work on the cerebellum with psycholinguistic and historical linguistic research. It finally provides the first experimental attempts towards the empirical validation of this synthesis, employing transcranial magnetic stimulation.

A neuroanatomical locus traditionally seen as limited to lower sensorimotor functions, the cerebellum has, over the last decades, emerged as a widely accepted foundation of feedforward control and state estimation. Its cytoarchitectural homogeneity and diverse connectivity with virtually all parts of the central nervous system strongly support the idea of a uniform, domain-general cerebellar computation. Its reciprocal connectivity with language-related cortical areas suggests that this uniform cerebellar computation is also applied in language processing. Insight into the latter, however, remains an elusive desideratum; instead, research on cerebellar language functions is predominantly involved in the frontal cortical-like deficits (e.g. aphasias) seldom induced by cerebellar impairment. At the same time, reflections on cerebellar computations in language processing remain at most speculative, given the lack of discourse between cerebellar neuroscientists and psycholinguists.

On the other hand, the fortunate contingency of the recent accommodation of these computations in psycholinguistic models provides the foundations for satisfying the desideratum above. The thesis thus formulates a neurolinguistic model whereby multi-level, predictive, associative linguistic operations are acquired and performed in neocerebello-cortical circuits, and are adaptively combined with cortico-cortical categorical processes. A broad range of psycholinguistic phenomena, involving, among others, ‘pragmatic normalization’, ‘verbal/semantic illusions’, associative priming, and

phoneme restoration, are discussed in the light of recent findings on neocerebellar cognitive functions, and provide a rich research agenda for the experimental validation of the proposal.

The hypothesis is then taken further, examining grammaticalization changes in the light of neocerebellar linguistic contributions. Despite a) the broad acceptance of routinization and automatization processes as the domain-general core of grammaticalization, b) the growing psycholinguistic research on routinized processing, and c) the evidence on neural circuits involved in automatization processes (crucially involving the cerebellum), interdisciplinary discourse remains strikingly poor. Based on the above, a synthesis is developed, whereby grammaticalization changes are introduced in routinized dialogical interaction as the result of maximized involvement of associative neocerebello-cortical processes.

The thesis then turns to the first steps taken towards the verification of the hypothesis at hand. In view of the large methodological limitations of clinical research on cerebellar cognitive functions, the transcranial magnetic stimulation apparatus is employed instead, producing the very first linguistic experiments involving cerebellar stimulation. Despite the considerable technical difficulties met, neocerebellar loci are shown to be selectively involved in formal- and semantic-associative computations, with far-reaching consequences for neurolinguistic models of sentence processing. In particular, stimulation of the neocerebellar vermis is found to selectively enhance formal-associative priming in native speakers of English, and to disrupt, rather selectively, semantic-categorical priming in native speakers of Modern Greek, as well as to disrupt the practice-induced facilitation in processing repeatedly associated letter strings. Finally, stimulation of the right neocerebellar Crus I is found to enhance, quite selectively, semantic-associative priming in native speakers of English, while stimulation of the right neocerebellar vermis is shown to disrupt semantic priming altogether. The results are finally discussed in the light of a future research agenda overcoming the technical limitations met here.

Declaration

I hereby declare that this thesis is of my own composition, and that it contains no material previously submitted for the award of any other degree or professional qualification. The work reported here has been executed by myself, except where due acknowledgement is made in the text. The thesis complies with all regulations for the degree of PhD at the University of Edinburgh and falls below the requisite word limit specified.

Giorgos P. Argyropoulos

August, 2010

Acknowledgements

I gratefully acknowledge the support of my supervisors, Dr. Thomas Bak, Dr. Patrick Sturt, and Prof. Jim Hurford in the preparation of the thesis and in funding applications.

Research in both transcranial magnetic stimulation (TMS) and cerebellar function remained, unfortunately, outwith the contexts of the Edinburgh academia. I would thus like to thank a number of people that provided significant support for the research reported here. I'm grateful to Dr. Richard Shillcock for encouraging me to explore cerebellar functions with TMS, and for initiating my 'assimilation' to the TMS 'borg'. Very little would have been possible without him. Special thanks goes to Dr. Neil Muggleton for hosting me in the Institute of Cognitive Neuroscience (University College London), providing his help and expertise on many technical issues of TMS, and recruiting the subjects. I am also grateful to Dr. Gorana Pobric (Neuroscience and Aphasia Research Unit, University of Manchester) for her support in my first experimental attempts with the TMS apparatus. Furthermore, I would like to thank Prof. Vasilios Kimiskidis and Dr. Sotiris Papagiannopoulos, for hosting my experiment in the 3rd Neurological Clinic of the 'Papanikolaou' Hospital in Thessaloniki, and for all their precious advice, support, and encouragement. Thanks goes to Prof. Chris Miall, for his assured responses to my questions on the cerebellum, as well as to Prof. John Desmond, Prof. Massimiliano Oliveri, and Prof. Giacomo Koch for all their precious advice on issues of cerebellar TMS; also, to Dr. Joe Devlin, Dr. Narender Ramnani, and Prof. Vince Walsh, for their interest and encouragement.

It was also a privilege to discuss a range of neuroscience-related issues with Prof. Terrence Deacon, Dr. Andrea Gebhart, Prof. Peter Hagoort, and Prof. Philip Lieberman. A number of psycholinguists also provided their input and took interest in my work: I would therefore like to thank Prof. Fernanda Ferreira, Dr. Manon Jones, Prof. Ken

McRae, Prof. Martin Pickering, Prof. Arthur Samuel, and Dr. Alice Turk for their help. The work here also received input and attention from some outstanding historical and theoretical linguists: Prof. Joan Bybee, Dr. Ronnie Cann, Prof. Peter Harder, Prof. Martin Haspelmath, and Prof. Steve Nicolle– I am profoundly grateful for their precious input. It is of course close to redundant to say that not all of the people above have to agree with what is said here, and none of them are responsible for any mistakes found.

I would also like to thank a number of people that have helped and encouraged me in ever so many different ways these years, especially Dr. Stavros Assimakopoulos, Prof. Dimitris Gounelas, Prof. Philip Kargopoulos, Dr. Elvira Masoura, Prof. Miriam Meyerhoff, Dr. Kamran Rastegar, Prof. Savas Tsohatzidis, Dr. Lambros Varelas, as well as the people of the 'Efxinos Leschi'. Also, Jenny-Mara Gumer, Katie Keltie, Shirley Kirkaldy, Magdalena Maszurzak, and Toni Noble, for their precious support in all administrative issues, and Michael Benett, Ziggy Campbell, and Eddie Dubourg for their ever so important technical support. I warmly acknowledge the help and encouragement of my relatives, the Fostiropoulos, Lavasas, Monas, and Skoufas families for their precious help and encouragement. My thanks also go to the participants, especially the TMS group, both Greek- and English- native speakers.

The research conducted throughout these four years was funded by 'the Pavlos and Elissavet Papagiannopoulou Foundation', the support of which allowed me to devote full time to my studies.

Finally, this dissertation, which will hopefully prove in the near future to be but the beginning of more organized research into the new fields that it now opens, is dedicated to Prof. Talmy Givón, whose work has strongly inspired me ever since I entered the discipline of Linguistics. Having my first publication cited with such kind words in his latest work (Givón, 2009) has been a great honor to say the least. The thesis is also dedicated to my professor, Tassos Christidis, whose Critical thought I dearly miss and owe a lot to.

Copyright Acknowledgement

The following figures, images, and publications were reproduced with permission from the corresponding publishing companies.

i. Figure 1.2 from 'The human cerebro-cerebellar system: Its computing, cognitive, and language skills', by H. C. Leiner, A. L. Leiner, & R. S. Dow (1991). *Behavioral Brain Research, 24,* 113-128. © 1991 Elsevier.

ii. Figure 1.3 from '*Human physiology: from cells to systems*' (7th ed.), by L. Sherwood (2010). Grove, California: Brooks/Cole. © 2010 Brooks/Cole.

iii. Figure 1.4 from 'Historical review of the significance of the cerebellum and the role of Purkinje cells in motor learning', by M. Ito (2002), *Annals of the New York Academy of Sciences, 978,* 273-288. © 2002 Wiley and Sons.

iv. Figure 1.5 from 'Expanding cerebellar horizons', by N. Ramnani & C. Miall (2001). *Trends in Cognitive Sciences, 5*(4), 135-136. © 2001 Elsevier.

v. Figure 2.1 from 'Movement and thought: Identical control mechanisms by the cerebellum', by M. Ito (1993). *Trends in Neurosciences, 16*(11), 448-450. © 1993 Elsevier.

vi. Figure 2.2 from 'Neural control of cognition and language', by M. Ito (2000). In A. Marantz, Y. Miyashita, & W. O'Neil (Eds.), *Image, language, brain* (pp. 149-162). © 2000 MIT Press.

vii. Figure 2.3 from 'Do people use language production to make predictions during comprehension?' by M. Pickering & S. Garrod (2007). *Trends in Cognitive Sciences, 11*(3), 105-110. © 2007 Elsevier.

viii. Figure 2.4 from 'Neocerebellar emulation in language processing', by G. P. Argyropoulos (2009). 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. No copyright owned by Media Tryck.

ix. Figure 2.5.A from 'On Broca, brain, and binding: a new framework', by P. Hagoort (2005). *Trends in Cognitive Sciences, 9,* 416-423. © 2005 Elsevier.

x. Figure 2.5.B from 'Control of mental activities by internal models in the cerebellum', by M. Ito (2008). *Nature Reviews Neuroscience, 9,* 304-313. © 2008 Nature.

xi. Figure 2.6 from 'Dysmetria of thought. Clinical consequences of cerebellar dysfunction in cognition and affect', by J. D. Schmahmann (1998). *Trends in Cognitive Sciences, 2,* 362-370. © 1998 Elsevier.

xii. Table 2.2 from ‘Constructions: a new theoretical approach to language’, by A. E. Goldberg (2003). *Trends in Cognitive Sciences*, *7*(5), 219-224. © 2003 Elsevier.

xiii. Figure 3.2 from ‘Is grammaticalization glossogenetic?’, by G. P. Argyropoulos (2010a). In A. D. M. Smith, M. Schouwstra, B. de Boer, & K. Smith (Eds.), *The Evolution of Language: Proceedings of the 8th International Conference on the Evolution of Language* (pp. 3-10). © 2010 World Scientific.

xiv. Figure 3.3 from ‘The subcortical foundations of grammaticalization’, by G. P. Argyropoulos (2008a). 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). © 2008 World Scientific.

xv. Figure 4.1 from ‘Neuroimaging studies of the cerebellum: language, learning, and memory’ by J. E. Desmond, & J. A. Fiez (1998). *Trends in cognitive sciences*, *2*(9), 355- 362. © 1998 Elsevier.

xvi. Figure 5.1 from ‘Size of the neocerebellar vermis is associated with cognition in healthy elderly men’ by A. M. J. MacLullich, C. L. Edmond, K. J. Ferguson, J. M. Wardlaw, J. M. Starr, J. R. Seckl, & I. J. Deary (2004). *Brain and Cognition*, *56*, 344-348. © 2004 Elsevier.

xvii. Figure 7.1 from ‘Cerebellar involvement in verb generation: An fMRI study’ by M. Frings, A. Dimitrova, C. F. Schorn, H.- G. Elles, C. Hein-Kropp, E. R. Gizewski, H. C. Diener, & D. Timmann (2006). *Neuroscience Letters*, *409*, 19-23. © 2006 Elsevier.

xviii. Publication: ‘Past and future, human and non-human, semantic/procedural and episodic’, by J. R. Hurford, M. Flaherty, & G. Argyropoulos (2007). *Behavioral and Brain Sciences*, *30*(3), 324-325. Commentary on T. Suddendorf & M. Corballis, *The evolution of foresight: What is mental time travel and is it unique to humans?* © 2007 Cambridge University Press.

xix. Publication: ‘The subcortical foundations of grammaticalization’, by G. P. Argyropoulos (2008). 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. © 2008 World Scientific.

xx. Publication: ‘Is grammaticalization glossogenetic?’, by G. P. Argyropoulos (2010). In A. D. M. Smith, M. Schouwstra, B. de Boer, & K. Smith (Eds.), *The evolution of language: Proceedings of the 8th international conference on the evolution of language* (pp. 3-10). Singapore: World Scientific. © 2010 World Scientific.

Abbreviations

A(P)	adjective, adjectival phrase
ACC	accusative case
ANOVA	analysis of variance
BA	Brodmann area
BNC	British national corpus
CB	cerebellum, cerebellar
CL	clitic (pronoun)
CNMC	corticonuclear microcomplex
DEM	demonstrative
DET	determiner
EAT	Edinburgh word association thesaurus
ERP	event-related potential
EZ	‘ezafe’ marker
(f)MRI	(functional) magnetic resonance imaging/image
GEN	genitive case
HNC	Hellenic national corpus
INF	infinitive
LAST	late assignment of syntax theory
LTD	long-term depression
MSe	mean square error
MEP	motor-evoked potential
ms	millisecond(s)
MUC	memory, unification, and control
NCB	neocerebellum, neocerebellar

NCBV	neocerebellar vermis
NCBKFLP	neocerebellar Kalman filter linguistic processor
NOM	nominative case
N(P)	noun, noun phrase
OBJ	object, object marker
PART	participle
PERF	perfect tense
PET	positron emission tomography
PLCB	posterolateral cerebellum/ cerebellar
PL	plural number
POSS	possessive pronoun
RC	relative clause
rCBF	regional cerebral blood flow
REL	relative pronoun
RT	reaction time(s)
S	sentence
SD	standard deviation
SEM	standard error of the mean
SG	singular number
SOA	stimulus onset asynchrony
(c)TBS	(continuous) theta-burst stimulation
(r)TMS	(repetitive) transcranial magnetic stimulation
tDCS	transcranial direct current stimulation
V(P)	verb (phrase)
θ-(role)	thematic (role), thematically
1, 2, 3	first, second, third person

Contents

LIST OF FIGURES

LIST OF PLOTS

LIST OF TABLES

Organization of the Thesis

The thesis is largely divided into two parts, **a theoretical** (chapters 1-3; Volume I), and an **empirical** one (chapters 4-7; Volume II), followed by a brief discussion of the findings and future directions (chapter 8).

In the **theoretical part** of the thesis **(Volume I), chapter 1** provides a necessary non-technical background to basic cerebellar (CB) anatomy and circuitry, contextualizing the synthesis and the model proposed in the next chapter. CB circuitry is strikingly homogeneous, prompting neuroscientists to consider a uniform CB computation.[1] The fact that different compartments reciprocally connect with different areas of the central nervous system indicates that such uniform computation is of a multimodal nature. The fact that fundamental language-related cortical areas show such reciprocal connectivity with the CB indicates that the same computation is extended to language processing. However, research on CB language pathology is preoccupied with the way in which CB impairments induce 'frontal-like' language deficits, e.g. due to cerebrocerebellar diaschisis or impaired modulation of cerebral linguistic functions. On the other hand, reflections on the nature of the involvement of this uniform, multimodal CB computation in language processing remain poor and speculative.

Expanding on Argyropoulos (2009), **chapter 2** presents the first model of neocerebellar (NCB) state estimation and feedforward control in language processing. The chapter picks up the discussion with the answer to the question left open at the end of Chapter 1, i.e. on the nature of CB computations. The thesis adopts the long-established and experimentally well-grounded theories that the CB is the storage space of internal models and/or the neuroanatomical analogue of a Kalman filter. In this context, the recent introduction of feedforward control and state estimation in psycholinguistic models represents a formidable opportunity to articulate concrete, computationally-based neurolinguistic hypotheses on the NCB. After a brief

[1] See main text for all references.

introduction to the necessary neurocomputational concepts, the chapter introduces the hypothesis of the Neocerebellar Kalman Filter Linguistic Processor (NCBKFLP; Argyropoulos, 2009), along with the idea that the NCB instantiates associative processes adaptively combined with cerebral, categorical ones. Subsequently, a number of sentence processing mechanisms are shown to rely on computations identical or strikingly similar to those of the NCB. These mechanisms are discussed in the light of supportive evidence from recent literature on NCB language functions. While covert imitation and mirror neuron functions remain beyond the scope of the thesis, the possibility is entertained that the framework developed here may be fully compatible with proposals on such mechanisms. The chapter then continues with the presentation of an experimental research agenda, explicitly based on psycholinguistic phenomena involving state estimation and feedforward control. The first elementary steps in implementing this agenda are made in chapters 5, 6, and 7.

In the line of Argyropoulos (2008a, b), **chapter 3** attempts an interdisciplinary synthesis among psycholinguistics, historical linguistics, and neuropsychology of language, by studying grammaticalization processes in the light of the NCBKFLP. The term 'grammaticalization' is used to describe a wide range of multi-level linguistic phenomena that occur during the historical change of lexical items into grammatical ones, giving birth to new syntactic constructions. First, the approach to grammaticalization phenomena is discussed. Following the line of reasoning in Argyropoulos (2010a), it is supported that such changes cannot be explained as by-products of structural misconvergences between transmission and reception. Instead, the thesis adopts a view particularly shared in the functionalist camp of grammaticalizationists,[2] where these changes are seen as products of routinized-automatized performance in intra-generational language transmission. Despite the psycholinguistic work on dialogical routinization and research on the circuits involved in automatization processes, interdisciplinary discourse remains strikingly poor. In view of this, the chapter goes on to briefly investigate the anatomical substrates of

[2] The term is often used in the relevant literature (e.g. Haspelmath, 2004).

automatization processes from a domain-general perspective, demonstrating the significance of CB circuitry. The discussion then turns to particular phenomena of grammaticalization processes, from phonetic attrition, cliticization, affixation, and univerbation, to pragmatic routines and mental backgrounding. The mechanisms employed by grammaticalizationists to describe such changes are shown to strongly reflect NCB, or, a fortiori, NCBKFLP computations. In particular, it is argued that grammaticalization is the larger-scale product of a shift from categorical to associative computations in processing particular instances of constructions in routinized dialogical contexts. As argued in chapter 2, these associative computations involve NCB circuitry, and such adaptive trade-offs are controlled by the NCBKFLP.

In the **empirical part** of the thesis **(Volume II)**, the first steps are taken to experimentally investigate the involvement of NCB state estimation and feedforward control in language processing. **Chapter 4** presents the principles of the paradigm adopted here, i.e. transcranial magnetic stimulation (TMS), in the context of research on CB cognitive functions. The reasons why TMS was preferred over a clinical paradigm are provided, elucidating how TMS overcomes many methodological obstacles in studying CB language functions. The chapter then continues with the discussion of the particular TMS apparatus and parameters employed here, in setting up the first linguistic CB TMS studies in the literature.

Chapter 5 reports and discusses the findings of the first TMS study undertaken (Argyropoulos, 2010b). Two different noun-to-noun priming phenomena were assessed before and after TMS of the right neocerebellar vermis (NCBV) and of a control, right posterolateral CB site in healthy native English speakers: 'formal-associative' priming (e.g. 'gift-horse') and 'semantic-categorical' priming (e.g. 'penny-coin'). TMS of the NCB vermis was found to selectively enhance formal-associative priming and induce some reading-related disruptions. The findings are discussed in the light of the NCBKFLP and of NCBV pathology in schizophrenic cognitive-linguistic impairments.

Chapter 6 reports and discusses the findings of the second TMS study (Argyropoulos *et al.,* 2010). Semantic-categorical priming was again contrasted with formal-associative priming before and after stimulation of the same CB sites in healthy native Modern Greek speakers. Despite a number of limiting conditions in stimulus construction and TMS application, a rather selective disruption of semantic-categorical priming and reading-related processes was found after NCBV TMS. Above all, there was a striking lack of decrease of reaction times after NCBV TMS that was strongly selective for the group that underwent NCBV TMS in the second session of participation. This pattern is discussed in the context of the significance of the NCB in processing repetitively co-occurring linguistic input.

Chapter 7 reports and discusses the findings of the third TMS study (Argyropoulos & Muggleton, 2010). Two distinct versions of semantic, noun-to-verb priming were assessed before and after TMS of the right NCBV and the right NCB Crus I in healthy native English speakers: categorical priming, where nouns and verbs are synonymous (e.g. 'robbery-stealing'), and associative priming, where the nouns are appropriate fillers of the thematic roles projected by the following verbs (e.g. 'broom-sweeping'). Despite the limitations in yielding priming effects in baseline conditions, TMS of the right Crus I was found to enhance semantic-associative priming quite selectively, while TMS of the right NCBV disrupted semantic priming overall. The findings are discussed in the context of NCBKFLP involvement in schema-transmission and mental model processing.

The thesis is concluded in **chapter 8**, with a discussion of future theoretical and experimental directions in the light of the findings and the limiting conditions in the research conducted so far. The discussion ends with the note that the foundations have been laid here to study NCB 'dynamics memory' in language processing, from both a synchronic and a diachronic perspective.

Volume I

Theoretical Chapters

Introduction to Volume I

The first part of the thesis (chapters 1-3) provides a synthesis among psycholinguistics, historical linguistics, and computational neuroscience in constructing a hypothesis on neocerebellar computations in language processing. A model is proposed, showing how neocerebellar feedforward control and state estimation are employed in language perception-comprehension. These processes are argued to provide the grounding for a number of fundamental psycholinguistic and historical linguistic phenomena.

Chapter 1

The Cerebellum and Language

> ‘‘[…] this is some terrific computer down here’’
> (Henrietta Leiner; cited in Blakeslee, 1994)

1.1. Introduction

This chapter provides a selective, non-technical introduction to cerebellar (CB) architecture, along with some recent findings on CB involvement in language processing, in order to contextualize the hypothesis promoted in chapter 2. For further details on CB neuroanatomy and function, readers may refer to manuals of general neuroscience and human physiology (e.g. Arbib *et al.*, 1998; Sherwood, 2010), or to introductory material on the CB (e.g. Miall, 2001; Barlow, 2002).

1.2. Cerebellar Architecture

The 'cerebellum' (Latin, 'little brain'), also referred to as the ''hind-brain'' (Holmes, 1939), is located in the posterior cranial fossa, astride the brainstem, beneath the occipital cortex and the 'tentorium cerebelli' (figure 1.1). While it occupies approximately 10% of total brain volume and accounts for 10-15% of brain weight, it is the largest structure in the human brain in terms of the number of neurons it contains (Williams & Herrup, 1988). This number exceeds that of the entire cerebral cortex[3] (Zagon *et al.*, 1977; Noback & Demarest, 1981); it amounts to about half of all the neurons of the nervous system; it is at least five times larger than that of other parts of the nervous system in humans, and corresponds to at least half the nerve cell population of the entire body in higher mammals (Arbib *et al.*, 1998, p. 265). The CB also occupies 40% of brain surface area. During phylogenetic evolution, it enlarged more dramatically than any other part of the brain except the cerebral cortex (Passingham, 1975; Brodal, 1981).

The CB is also one of the most widely connected structures, having physiological connections with all major divisions of the central nervous system (e.g. Schmahmann & Pandya, 1989; Middleton & Strick, 1994; Schmahmann, 1996; section 1.3.2 below). These facts, along with its high afferent-to-efferent axon ratio (40:1; Carpenter, 1991), suggest a massively integrative CB function. The massive interconnectivity of the CB is as striking as its ''essentially uniform, monotonously repetitive architecture''– in other words, ''there are no ''Brodmann areas'' in the cerebellum'' (Schmahmann, 2000, p. 206; section 1.2.3 below). Its cytoarchitectural homogeneity and massive interconnectivity have supported the idea that it applies a unitary computation across different behavioral modalities (e.g. Bloedel, 1992; Wolpert *et al.*, 1998; Schmahmann, 2000).

[3] Throughout the thesis, the terms 'cortical'/'cerebral', and 'cerebral cortex'/'cerebrum' will be used interchangeably, as no particular distinctions will be necessary for the argument. The term 'cerebellar (CB) cortex/-ical' will be used to refer to the cortical layers of the CB, thus clearly distinguishing it from the 'cerebral cortex'.

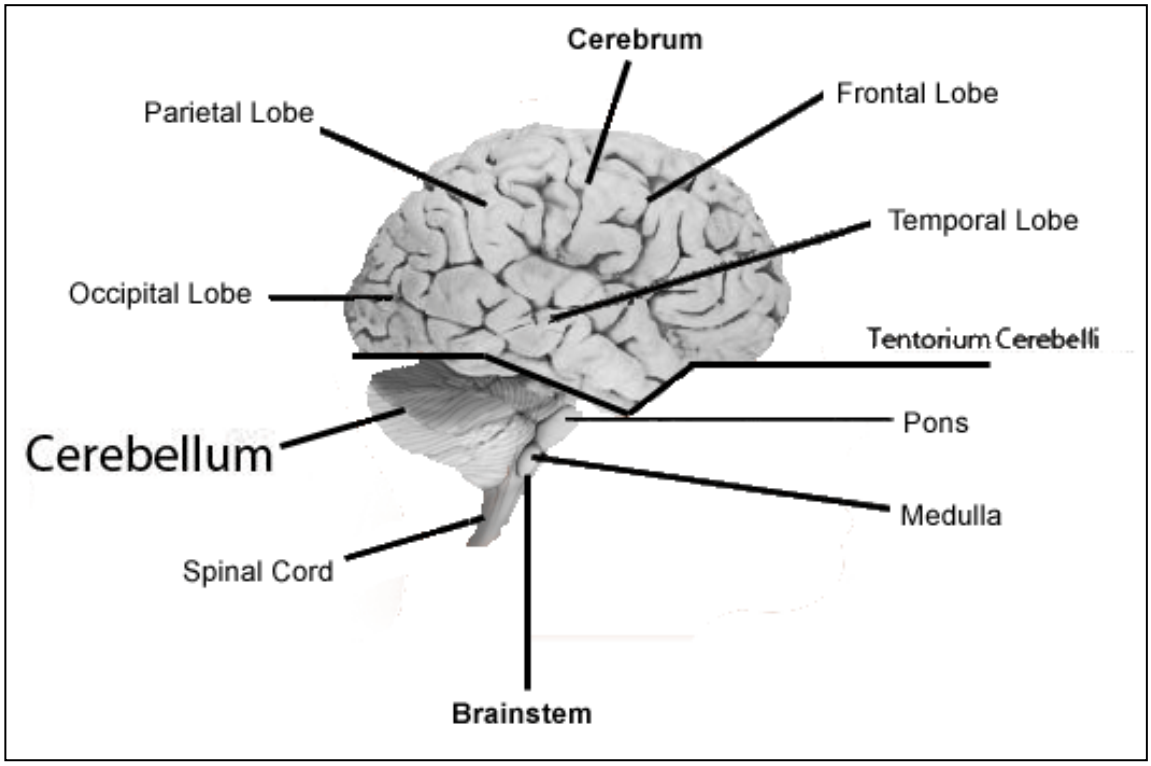


Figure 1.1: Location of the CB with respect to other major subdivisions of the human brain.

1.2.1. Major Cerebellar Subdivisions

The CB is divided into the 'flocculonodular lobe' and the 'corpus cerebelli', the latter consisting of two bilateral hemispheres connected by a narrow median portion, the 'vermis'. The corpus is subdivided into three broad longitudinal 'regions' or 'zones': 'medial' or 'vermal'; 'paramedial', 'paravermal', or 'intermediate'; and 'lateral' or 'hemispheric'. Each zone consists of an outer, surface gray mantle, the '(CB) cortex', a medullary core of underlying white matter composed of nerve fibers projecting to and from the CB, and a number of 'deep CB nuclei'.

The CB cortex is characterized by the homogeneous appearance of many long, transverse convolutions, generally oriented transversely and separated by ‘sulci’ or ‘fissures’. Two deep fissures, the ‘primary’ and ‘posterolateral’ ones, divide the CB into three lobes: ‘anterior’ and ‘posterior’ within the corpus, and ‘flocculonodular’ outwith the corpus. The primary fissure separates the anterior and posterior lobes of the CB. The anterior lobe is everything rostral to the primary fissure, while the posterior lobe is everything between the primary fissure and the posterolateral fissure. Each lobe consists of a vermal and a hemispheric component. Shallower fissures divide the anterior and posterior lobes into ten lobules (anatomists consider the flocculonodular lobe as the tenth lobule) in the vermis (I-X) and in the hemispheres (HI-HX). Accordingly, lobules I-V are part of the anterior lobe, and VI-IX are part of the posterior lobe (Larsell, 1937; figure 1.2).

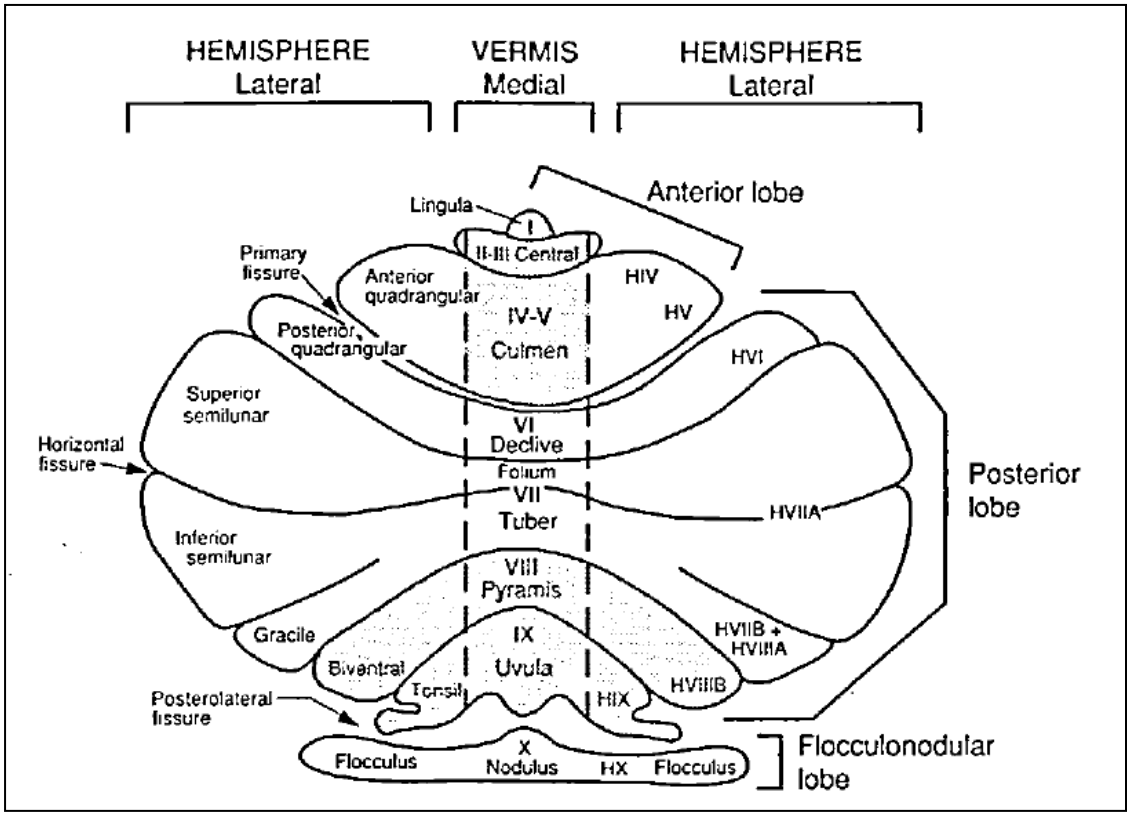


Figure 1.2: CB lobes and lobules. Figure reproduced from Leiner *et al.* (1991, p. 114) with permission © 1991 Elsevier.

In terms of its phylogeny and the different outputs of its compartments, the CB can be similarly subdivided into three main parts: the ‘archicerebellum’ outwith the corpus, and the ‘paleocerebellum’ and ‘neocerebellum’ within the corpus (figure 1.3).

a) The ‘archicerebellum’, ‘vestibulocerebellum’, ‘flocculonodular lobe’, or ‘flocculonodulus’ is the phylogenetically oldest part. It lies behind the posterolateral fissure, which separates it from the posterior lobe. It has no corresponding deep CB nucleus; it receives directly from the vestibular nuclei in the brainstem, and its output, from Purkinje cells of the CB cortex, projects back directly to them. In close connection with the vestibular system, it controls eye movements in relation to body position and motion (vestibular reflexes), and regulates balance, posture and trunk musculature.

b) The phylogenetically next oldest part is the ‘paleocerebellum’ or ‘spinocerebellum’. It is of vertebrate phylogeny, and comprises the anterior lobe and the medial and intermediate parts of the CB hemispheres. It receives spinal proprioceptive as well as auditory and visual inputs from the periphery and from the primary motor and somatosensory cortex. It projects information back to the spinal cord via the red nucleus, providing information about the progress of ongoing movements and correcting errors. It is concerned with orientation and precise control of proximal muscles. In general, the flocculonodular lobe, (anterior) vermis and paravermis are connected to the brain stem and the spinal cord, and confer adaptiveness on reflexes and compound movements, enabling animals to survive in ever-changing environments.

c) The ‘neocerebellum’ (NCB), ‘cerebrocerebellum’, or ‘pontocerebellum’ is the phylogenetically most recent part. It comprises the large posterior lobe, i.e. the posterior vermis, paravermis, and lateral hemispheres. It has undergone tremendous expansion in primate evolution, and is by far the largest CB subdivision in humans. It is involved in the regulation, planning, and learning of highly skilled, spatially and temporally complex movements, as well as in higher cognitive and affective processes (section 1.3). The NCB is connected, among other areas, to the motor cortex and is thus associated

with movements of limbs, feet, hands, and digits. Damage to intermediate and lateral CB parts disrupts ipsilateral arm, hand, and finger movements significantly more than trunk movement.

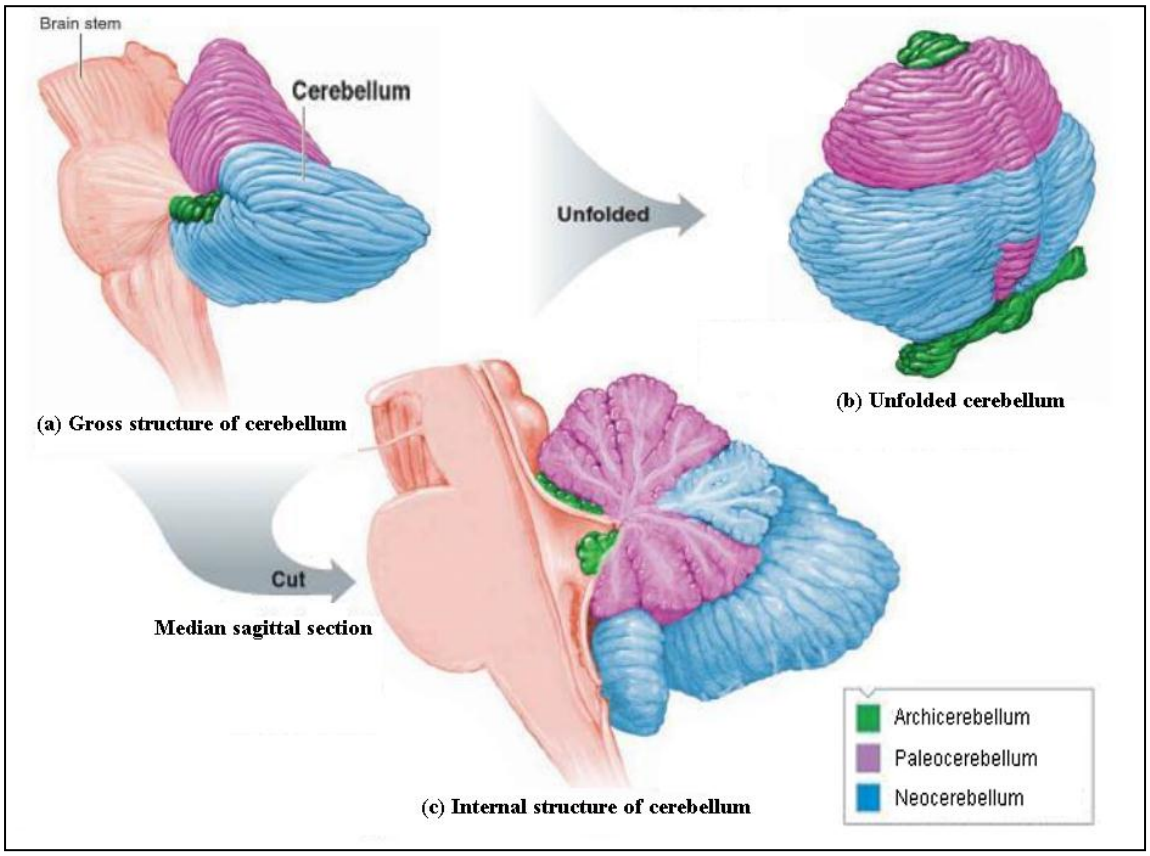


Figure 1.3: Phylogenetically-based division of the CB. Figure adapted from Sherwood (2010, p. 166) with permission © 2010 Brooks/Cole.

1.2.2. Deep Cerebellar Nuclei

Each CB cortical region projects in a topographically systematic fashion to the underlying deep nuclei, which in turn project, via the thalamus, to the cortex and brainstem, thus providing the output fibers of the CB. Hence, virtually the entire CB output is conveyed to the rest of the central nervous system via the projections of the CB cortex to the CB nuclei.[4] There are four major pairs (right and left) of deep nuclei: the dentate nuclei, the two pairs of interposed nuclei (globose and emboliform), and the fastigial nuclei. Each pair receives input from a different region of the CB cortex. In general, the NCB projects primarily to the dentate nucleus, the paleocerebellum (paravermis in particular) to the interposed nuclei, and the archicerebellum to the fastigial nucleus.

1.2.3. Cerebellar Cortical Layers and Circuitry

Despite the overwhelming number of its neurons, the CB contains relatively few neural cell types: the inhibitory stellate, basket, Purkinje, and Golgi neurons, and the excitatory granule cells. These five main neurons are organized in three layers in the CB cortex:

a) The **outermost** or **'molecular layer'** contains the bodies of stellate and basket cells, dispersed among the excitatory granule cell axons and inhibitory Purkinje cell dendrites; b) Beneath the molecular layer lies the thin middle or **'Purkinje cell layer'**, defined by Purkinje cell bodies; c) The **innermost** or **'granular layer'** contains a huge number (approximately 10^{11}) of granule cells and a few larger Golgi interneurons.

[4] Except for a direct projection of vestibulocerebellar Purkinje cells to one of the vestibular nuclei (the 'Deiters nucleus').

1.2.3.1. Purkinje Cells

The ultimate destination of the afferent pathways to the CB cortex is the Purkinje cells in the middle CB layer. These neurons have large cell bodies and fan-like dendritic arborizations extending upward into the molecular layer. They are the only output cells of the CB cortex, with their axons projecting into the underlying white matter to the deep CB or vestibular nuclei. Their output is mediated by the neurotransmitter g-aminobutyric acid and is thus wholly inhibitory. However, the deep CB nuclei receive excitatory input from the collaterals of the mossy (and climbing) fibers that ascend to the CB cortex to reach granule cells (see below). Thus, Purkinje cell inhibition of the deep nuclear neurons serves to modulate the level of such excitation. Depending on the prevalence of either nuclear or cortical activity, the CB input to the cerebrum may modulate tonically, enhancing excitability, or phasically, reducing the excitability of cortical outputs (Molinari *et al.*, 2002).

Purkinje cells receive inhibitory input by stellate, basket, and Golgi cells,[5] and excitatory input from two afferent fiber systems: the climbing fiber, and the granule cell-mossy-parallel fiber system. Both climbing and mossy fibers convey excitatory input directly from the spinal cord and brainstem to both deep CB nuclei and the CB cortex.

1.2.3.2. Mossy Fiber Input

Neurons in the pontine nuclei receive a projection from the cerebrum and relay such information to the contralateral CB cortex. The axons from the pontine nuclei, called ‘mossy fibers’, synapse on granule cells in the CB cortex. The CB granule cells give rise to specialized axons called ‘parallel fibers’, ascending to the molecular layer of the CB cortex. Parallel fibers carry some combination of activity on several mossy fibers

[5] Those three cell types will receive no emphasis in the subsequent chapters, unlike mossy and climbing fibers.

to the molecular layer, relaying information via excitatory synapses onto the dendritic spines of the Purkinje cells. This parallel fiber input creates the 'context' that each Purkinje cell may recognize. Each Purkinje cell is in a position to receive input from nearly a million parallel fibers, and each parallel fiber can contact tens of thousands of Purkinje cells. Mossy fibers thus indirectly excite the dendritic trees of Purkinje cells via the granule cell dendrites as well as their axonal branches in the form of parallel fibers.

1.2.3.3. Climbing Fiber Input

Purkinje cells also receive direct modulatory input from climbing fibers**,** all arising in the contralateral inferior olivary nucleus in the medulla. Every climbing fiber divides within the granular layer into branches, each of which enters the molecular layer and makes up to several hundred synaptic contacts on the adjacent dendrites of a single Purkinje cell.

Climbing fibers induce selective 'long-term depression' (LTD) in the synaptic strength of parallel fibers that are concurrently active, producing long-lasting effects on their synaptic efficacy (Ito *et al.*, 1982). Their activity may thus create and change the mossy fiber-parallel fiber 'activity context' driving a Purkinje cell. In the acquisition or adaptation of a movement, the climbing fiber, normally firing irregularly at a rate of approximately 1 Hz, suddenly, driven by errors, e.g. between the intended and the actual movement, fires once, immediately after the error, and reliably time after time. The effect of this is to reduce the strength of the synapse on the Purkinje cell of the parallel fibers active at that time, e.g. those participating in an inappropriate movement. Consequently, after practice and repeated firings, only the parallel fibers involved in the correct movement are left. Upon elimination of the error, the climbing fiber returns to its random background firing (Thach, 1997).

1.2.4. The Cerebellar Corticonuclear Microcomplex

The CB ‘‘corticonuclear microcomplex’’ (Ito, 1984; henceforth, CNMC) crystallizes this repetitive CB microcircuitry described above into a ‘‘modular unit of the cerebellum’’ (Ito, 2002, p. 281). A CNMC consists of a ‘‘microzone’’, i.e. a small area of the CB cortex, and a small number of nuclear cells. As shown in figure 1.4 below, the microzone receives two basic kinds of input, mossy and climbing fibers, and the output is carried by the deep CB nuclei. The set of mossy fiber afferents descending from pre-cerebellar loci conveys the major input to a CNMC, and provides excitatory synapses via collaterals to the nuclear neurons. The signals of the latter are under inhibitory regulation by the Purkinje cells of the CB cortex, which are the only output cells of the microzone. The mossy fiber input is at the same time transformed by the granule cells, whose axons form the parallel fibers, which in turn provide excitatory synapses to Purkinje cells. Climbing fiber inputs convey error signals to the CNMC, and LTD occurs at parallel fiber-to-Purkinje cell synapses after conjunctive activation of these synapses together with climbing fiber-to-Purkinje cell synapses. Conceivably, in the initial stage of practicing an activity, all parallel fiber synapses on Purkinje cells may be functional, so that Purkinje cells are fully activated by a mossy-fiber input. If the consequent silence of nuclear neurons results in performance errors, climbing fiber signals are transmitted, depressing parallel fiber-to-Purkinje cell signalling, and thereby releasing nuclear neurons from Purkinje cell inhibition and enabling CNMC output signals. In this way, CNMCs would be able to adaptively modify their performance toward the minimization of errors (Ito, 2002, pp. 281-2).

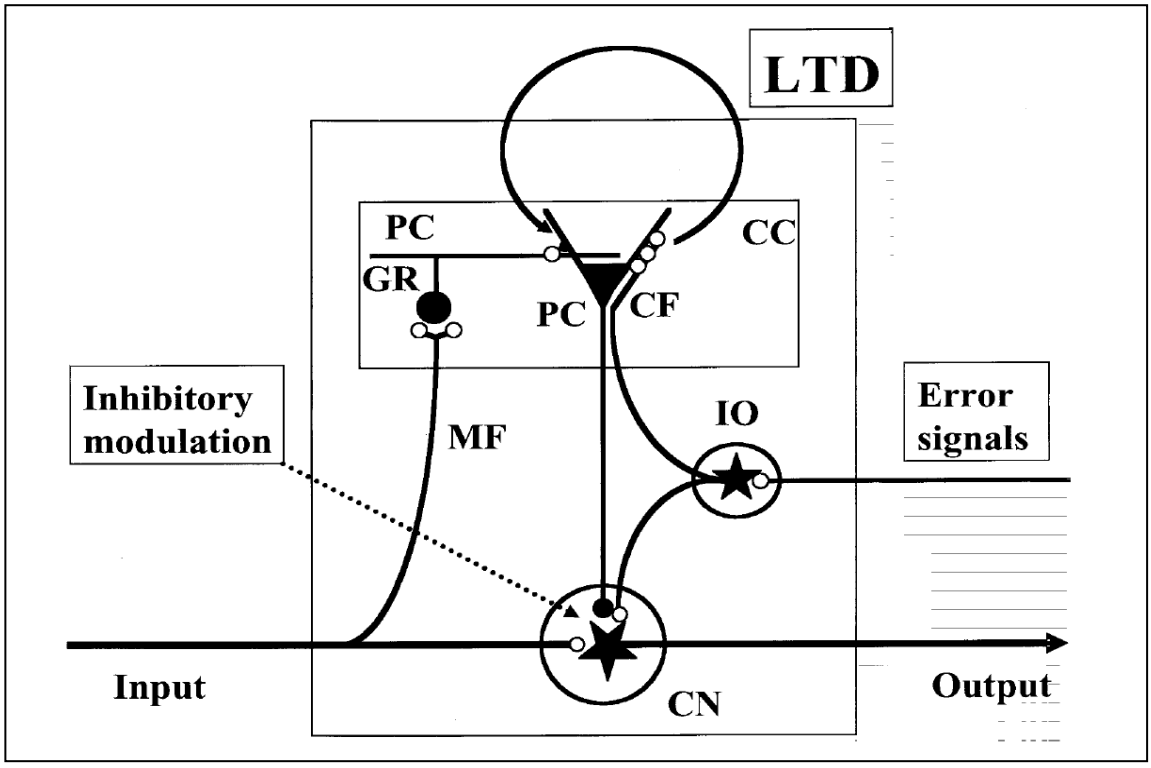


Figure 1.4: The CNMC (Ito, 1984). Long-term depression (LTD) occurs at parallel fiber-to-Purkinje cell synapses after their conjunctive activation with climbing fiber-to-Purkinje cell synapses (Ito *et al.*, 1982). PC: Purkinje Cell; GR: Granule Cell; CC: CB Cortex; IO: Climbing Fiber; MF: Mossy Fiber; CN: CB Nucleus; CF: Climbing Fiber; LTD: Long-Term Depression; hollow circle: excitatory synapse; filled circle: inhibitory synapse. Figure reproduced from Ito (2002, p. 281) with permission © 2002 Wiley and Sons.

1.3. Language Functions

1.3.1. Historical Background

Being ‘‘the long-time favorite brain structure of motor learning theorists’’ (Hazeltine & Ivry, 2002, p. 1979), the CB was traditionally seen as exclusively involved in motor functions, e.g. in the coordination of skilled, voluntary movement, gait, posture, balance, control of muscle tone, motor learning and articulation (Ivry & Fiez, 2000). As recorded in the monograph of Dow and Moruzzi (1958), such involvement has been reported for nearly two centuries. Pioneers in CB research in the nineteenth century observed symptoms arising from CB lesions in animals and humans. For instance, Luigi Rolando discovered that CB lesions resulted in motor disturbances. Marie-Jean-Pierre Flourens observed that animals with CB lesions showed movement with spared spontaneity but lacking coordination. Luciani examined recovery after CB resection. Joseph Babinski recognized dysmetria (see section 2.2.2 below) as a unique characteristic sign of CB diseases. During World War I, Holmes reported that soldiers with gunshot wounds through the CB exhibited significantly slower onset of arm movement on the damaged side. He also described a pattern of abnormal and scanning speech with indistinct articulation in CB patients, which is nowadays established as a feature of ataxic dysarthria (e.g. Hartelius *et al.*, 2000; section 2.2.2 here). From this perspective, then, CB patients would be expected to exclusively exhibit motor-related speech impairments, with spared cognitive aspects of linguistic abilities.

1.3.2. Recent Findings

However, much as in the case of the basal ganglia (e.g. Middleton & Strick, 1994, 2000), anatomical, clinical, and, recently, neuroimaging (PET and fMRI) studies provide cumulative evidence for CB involvement in a wide range of higher cognitive functions, such as memory, executive functions, visuospatial processing, emotional regulation, thought modulation, and, crucially, language.

The initial claims for such contributions largely relied on anatomical data concerning the phylogenetically recent reciprocal cortico-cerebellar projections (e.g. Brodal, 1981; Ito, 1984). The architectural heterogeneity of the dentate nucleus had been quite early recognized, with Dow (1942) defining a phylogenetically older, dorsomedial part, with minimal gyration and large neurons, and a recent, ventrolateral, heavily folded one, the 'neodentate', containing small neurons. In their seminal theoretical paper, Leiner *et al.* (1986; elaborated in Leiner *et al.*, 1989, 1991, 1995) proposed that, while this newer part was present in lower primates, only in anthropoids and humans had it evolved into a distinct region. They further emphasized that its massive expansion in humans had paralleled not only the expansion of NCB hemispheres, but also that of cerebral association areas (prefrontal cortex in particular). They also proposed that the evolution of the neodentate was paralleled by the differentiation of a ventrolateral thalamic nuclear area, through which the human neodentate would target the prefrontal cortex. This would involve the NCB in ''mental skills'' (*ibid*).

The recent application of retrograde and anterograde tracing studies[6] has corroborated these initial considerations. The cerebello-cerebral loop is nowadays established to be organized in multiple, topographically precise, segregated parallel channels, with each cerebral cortical area sending efferents to the CB via the pontine nuclei, while also being the recipient of a returning channel via the cerebello-thalamic

[6] 'Retrograde': tracing neural connections from the synapse to the cell body; 'anterograde': tracing neural connections from the cell body to the synapse.

tract. In particular, retrograde tracing studies have demonstrated that, while the elder part of the dentate sends projections to motor cortical areas, the neodentate projects, via the thalamus, not only to motor, but also to a range of 'cognitive areas'; among others, these include Brodmann areas (BA) 9/46 (dorsolateral prefrontal cortex) and 44/45 (Broca's area), as well as temporal areas, especially the superior temporal sulcus, and posterior parietal association areas (Middleton & Strick, 1994, 1997, 2001; Dum & Strick, 2003). On the other hand, anterograde tracing studies have shown that such cortical association areas send efferents to the CB cortex, via the basilar pons (Schmahmann & Pandya, 1995, 1997; Schmahmann, 2001). These include the superior and inferior parietal lobules, the temporal lobe, the superior temporal sulcus (including Wernicke's area), and prefrontal cortical areas; most importantly, BA 8A (control of conjugate eye movements), BA 9/46 (working memory processes), BA 10 (planning of actions, foresight, and judgment tasks), BA 32 and 12 (motivational and decision-making capabilities), and, above all, areas homologous to the language areas in humans (BA 44,45; e.g. Petrides & Pandya, 1994; see Schmahmann, 2001, p. 250, and Fabbro, 2000, for references). In addition to projections from the cerebral cortex, the phylogenetically older limbic lobe (reticular nuclei in the brainstem) and hypothalamus (BA 23, 24, 25 and 35) also project to the pontine nuclei, thus providing the CB with information for regulating autonomic, motivational and emotional behavior. Output fibers to those areas allow the CB to transmit information back to the limbic structures (Fabbro, 2000; Schmahmann, 2001).

This pattern is probably reproduced throughout the cerebro-cerebellar system (Schmahmann & Pandya, 1997, p. 54), and has supported the idea that the aforementioned unitary CB computations (section 1.2) are applied to a range of different processes, determined, each time, by the cerebral compartment interacting with the corresponding CB lobule (e.g. Schmahmann, 2000). An illustration of such loops is given in figure 1.5. The reciprocal connections of the CB with the cerebrum have thus added weight to the view that

> […] the cerebellum modifies behaviorally relevant information that it has received from the cerebral cortex via the corticopontine pathway and then redistributes this now "cerebellar-processed" information back to the cerebral hemispheres (Schmahmann & Pandya, 1997, p. 32).

Indeed, the discovery of such bisynaptic, reciprocal neural pathways between the NCB and language-related cerebral cortical areas (predominantly Broca's area, supplementary motor area, posterior parietal cortex and superior temporal sulcus) has provided the foundation for arguing for NCB involvement in non-motor language processes. Neuroimaging and clinical evidence corroborate this picture. Increased CB blood flow was initially found in a PET study, during covert word and rhyme generation in response to lexical stimuli (Petersen *et al.*, 1989). Similar findings of concomitant cortical (Broca's area and supplementary motor area) and CB activations have been more recently reported in such tasks, as well as in reading, word-stem completion, and naming actions in response to line drawings (for references and discussion, see Fiez & Raichle, 1997; Fabbro, 2000; Fabbro *et al.*, 2004). Similarly, in the clinical field, a case study on a right NCB patient showed very particular linguistic deficits: despite their excellent performance in memory, intelligence, and language function batteries, they showed deficits in practice-related learning and error detection (Fiez *et al.*, 1992; sections 2.4.1 and 2.4.2 here).

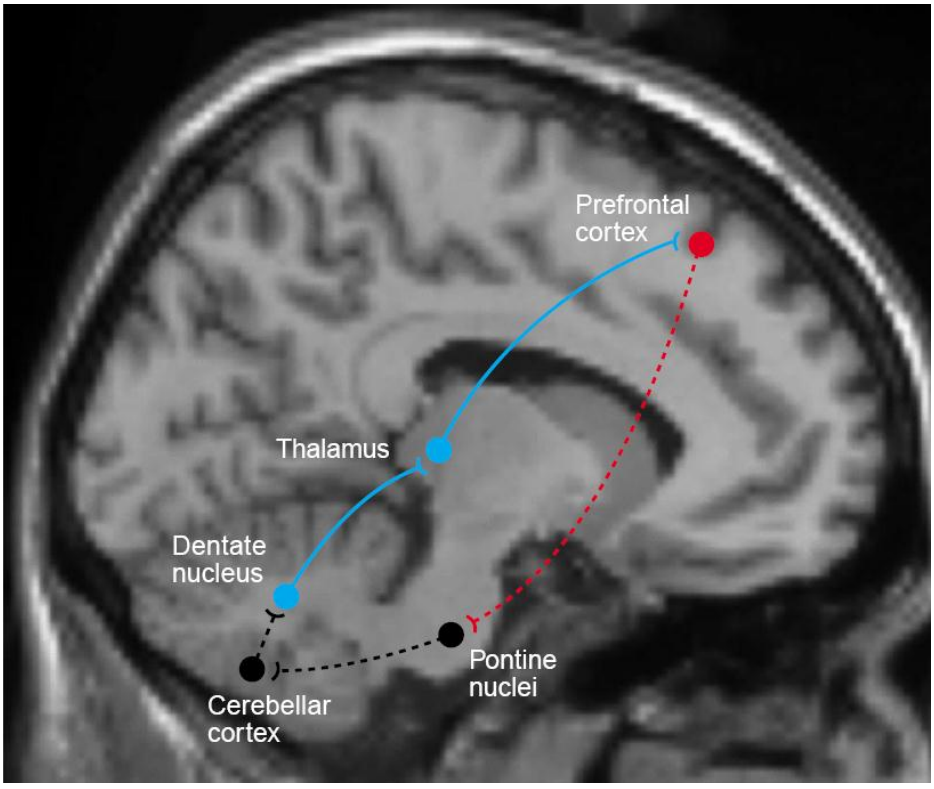


Figure 1.5: Cerebrocerebellar reciprocal connectivity. Blue: Projections from the NCB cortex to the dentate nucleus, and, via the thalamus, to prefrontal cortical areas. Red: Projections from the prefrontal cortex via the basilar pons back to the NCB cortex. Figure reproduced from Ramnani & Miall (2001, p. 135) with permission © 2001 Elsevier.

1.3.3. Beyond the Cerebral Chauvinism

Much like the basal ganglia, then, the CB has been shown to participate in a range of behavioral modalities far broader than any traditional motor-based account would acknowledge. This is why these subcortical areas are nowadays held to have domain-general contributions with particular computational properties, which can be identified on the grounds of their well-studied motoric functions:

> […] both the CB and the B[asal]G[anglia] exhibit massive interconnectivity with various areas of the cortex. Their interaction with these areas is probably not necessarily motor-based, but rather dependent on the nature of the process in each cortical area. The strategy of considering the kind of computations which are carried out by the CB and B[asal]G[anglia] in motor processing may reveal the nature of their contribution to other cognitive functions (Stowe *et al.*, 2004, p. 294).

However, despite the cumulating clinical evidence for even a ‘‘lateralized linguistic cerebellum’’[7] (Mariën *et al.*, 2001), little, if anything, has been said about the involvement of those unitary, multimodal computations (see section 1.2 above) in language processing. Instead, the study of CB involvement is constrained by the latent assumption that CB language deficits, if any, should reflect ‘cerebral-like’ ones. Characteristically, on the basis of anatomical and clinical evidence, it has been advocated for a long time that CB patients might exhibit a ‘‘mild frontal-like syndrome’’ (Botez *et al.*, 1985; Grafman *et al.*, 1992). For instance, CB patients have shown impairments on verbal, spatial, and intelligence measures in batteries of standard neuropsychological tests (Botez *et al.*, 1989; Bracke-Tolkmitt *et al.*, 1989).

Similarly, a number of cases of productive and receptive CB-induced agrammatism have followed up (see Mariën *et al.*, 2001, for a review). Silveri *et al.* (1994) first documented an association between right CB focal damage and transient expressive agrammatism. Syntactic symptoms typically involve lesions in right CB compartments (Riva & Giorgi, 2000; although see Fabbro *et al.*, 2000, for some exceptions), and frequently show satisfactory recovery. Morphosyntactic production errors, decreased utterance length, and erroneous comprehension of demanding syntactic structures are characteristic features. The explanation often promoted is that

[7] The degree of laterality of cerebellar linguistic functions is beyond the scope of the investigation here. Emphasis is given here to the right CB, which reciprocally connects with the left cortical loci.

agrammatism results as an adaptation to a deficit of a peripheral mechanism, e.g. working memory (Silveri *et al.*, 1994; Molinari *et al.*, 1997). In such an approach, then, the CB plays a modulatory role for the supratentorially represented cognitive and linguistic functions (Silveri & Misciagna, 2000), e.g. ensuring their correct execution (Daum & Ackermann, 1997). [8] Other cases of CB-induced agrammatism may even be explained on the basis of ‘‘crossed cerebro-cerebellar diaschisis’’, where (right) CB damage is followed by contralateral (left) cerebral cortical hypoperfusion (Broich *et al.*, 1987). In these cases, linguistic deficits are caused by functional depression of supratentorial language areas due to reduced cerebello-cortical input (Mariën *et al.*, 2001). Similarly, in speech pathology, CB ataxic dysarthria (section 1.3.1 here) exhibits commonalities with the articulation disorders in anterior perisylvian lesions, e.g. apraxia of speech and Broca’s aphasia, suggesting close cooperation between these areas and the CB (e.g. Mariën *et al.*, 2000). CB mutism is similarly explained by the fact that all these structures are traversed by pathways connecting the CB with supratentorial, language-related cerebral cortical areas, e.g. pre-motor areas, supplementary motor areas and Broca’s area (see Silveri & Misciagna, 2000; Justus & Ivry, 2001, pp. 276-7).

Such emphasis on the frontal-like identity of CB language pathology is justifiable from a methodological perspective. The early availability of psychological batteries for the pathology of the frontal cortex, often celebrated as the ‘‘cortex of cognition’’ (Goldman-Rakic, 1995), the very recent growth of CB cognitive pathology, as well as the subtle, possibly non-fundamental CB contributions to language, make such frontal-like CB impairments more easily accessible for empirical investigation than any CB computation-based accounts. However, commitment to a frontal-like perspective overlooks the very aspects of CB linguistic deficits that would instead directly reflect CB computations (see section 2.2). In such fashion, the gap yawning between research on the CB neurocomputational properties and CB language functions remains hard to bridge.

[8] Syntactic deficits represent the exception, rather than the rule, in CB pathology (Daum *et al.*, 2001); see chapter 2.

1.4. Conclusion: From ‘What’ to ‘How’

In this chapter, the CB was shown to have ‘‘a unitary principle of architecture, but a localizationist organizing principle of the connectivity’’ (Schmahmann, 2000, p. 206), supporting a unitary CB computation on multimodal input information. This ‘‘Universal Cerebellar Transform’’, or ‘‘cerebellumizing’’ (*ibid*, p. 207) may equally pertain to arousal-related, autonomic, affective, sensorimotor, cognitive and linguistic functions in a ‘‘meta-systemic’’ fashion (MacKay & Murphy, 1979; Xiang *et al.*, 2003). These properties shift attention from ‘what’ the CB computes to ‘how’ it does so (Hazeltine & Ivry, 2002, p. 1980).

However, what is the nature of this multimodal, unitary, metasystemic CB computation, and, in particular, how could it apply to language processing? Chapter 2 will begin with the introduction of a widely adopted framework of CB computations, and will attempt a synthesis with psycholinguistic models and phenomena.

Chapter 2

The Neocerebellar Kalman Filter Linguistic Processor

> ‘‘It is not clear how to explain recently observed cerebellar involvement in cognitive tasks such as language […] using the state estimator hypothesis (independent of associated motor behaviors or silent rehearsal of those behaviors), in large measure because the computational mechanisms underlying such tasks remain unclear. However, the hypothesis predicts that cognitive processes found to involve the cerebellum will be found to involve state estimation’’ (Paulin, 1997, p. 528).

2.1. Introduction

Chapter 1 concluded the discussion by leaving two questions open: what is the identity of this unitary, domain-general, metasystemic computation of the cerebellum (CB)? And how could such computation be employed in language processing? The present chapter briefly presents a widely adopted neurocomputational theory of the CB (section 2.2), and carries on to synthesize it with psycholinguistic models of language perception and comprehension (section 2.3). Clinical and imaging evidence on the CB is reviewed in the light of this synthesis, along with a number of psycholinguistic processes shown to directly rely on computations supported by NCB circuitry (sections 2.4 and 2.5). The chapter closes with certain experimental hypotheses directly derived from the model (section 2.6).

2.2. Cerebellar Feedforward Control and State Estimation

Despite its well-understood structure, the CB is remarkable for the number of conflicting theories put forward to account for its role (Miall, 2001). Treatment of those is far beyond the scope of the work here. Rather, the thesis will adopt an empirically assessed and widely accepted view that the CB functions as ‘‘the brain’s fortune teller’’ (Molinari *et al.*, 2009, p. 401), or as the ‘‘brain’s virtual reality machine’’ (Miall, 2007). In other words, the CB ‘‘predicts’’ and ‘‘prepares’’ the internal conditions required for a particular operation, be it sensory, motor, autonomic, memory-related, attention-related, affective, or linguistic (Courchesne & Allen, 1997). It does so by acquiring the ‘‘[…] predictive relationships among temporally ordered multidimensional sequences of exogenously derived […] and endogenously derived […] neural activities’’ (*ibid*, p. 2).

These predictive processes may be executed by altering neural response thresholds or cerebral blood flow levels in systems expected to participate in upcoming events, and by enhancing signals against noise. The consequences of CB output are fed back to it, modifying its associations between events and predictions (Courchesne & Allen, 1997). This ‘‘preparatory learning’’ (Courchesne & Allen, 1997) consolidates a ‘‘dynamics memory’’, which is reduced neither to procedural, nor to declarative aspects, but may optimize processes in both modalities (Ito, 1993 a, b).

From a neurocomputational standpoint, CB microcircuitry functions as such by storing ‘‘internal’’ (Ito, 1984), or ‘‘forward’’ (Kawato *et al.*, 1987) models of dynamic systems (see also Kawato & Gomi, 1992; Wolpert *et al.*, 1998), and/or by performing computations of state estimation (e.g. Paulin, 1989, 1993, 1997; Miall & King, 2008).

2.2.1. Feedforward Control and State Estimation

A brief, non-technical introduction to feedforward control and state estimation is now in order.

2.2.1.1. Feedforward Control

In supervised learning, the algorithm is presented with training data consisting of examples that include both the inputs and the desired outputs. The target output may be provided by an external teacher, or it may be specified internally. The latter case involves the acquisition of such an internal forward model, which allows the prediction of the consequences of actions made for the accomplishment of a goal. This model is then used to transform the input about these consequences into information for making the necessary internal modifications (Jordan & Rumelhart, 1992).

In control theory-parlance, internal models are systems mimicking the input-output transformations of the dynamic behavior of controlled objects (Jordan & Wolpert, 2000). Controlled objects convert a command issued by a controller into an output action. In order to mimic these transformations, internal models use a copy of the command issued to the controlled object to predict its consequences.[9] Predicted and actual consequences are then compared, and the discrepancy is fed back to the internal model to train it. In the context of the central nervous system, internal models may mimic the input-output characteristics of the motor apparatus. For instance, a 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 ('efference') issued by the central nervous system, and outputs a prediction of the future position and velocity of the arm (Jordan & Wolpert, 2000).

[9] While forward internal models predict consequences from copies of issued commands, ''inverse internal models'' calculate necessary feedforward commands from desired trajectory information (Kawato, 1999; Jordan & Wolpert, 2000). For the remainder of the thesis, the discussion will revolve around the former.

The significance of internal models is best illustrated in biological contexts, where the central nervous system is not immediately updated on changes in the peripheral motor system ('afference'), and any recent motor commands issued may be yet to affect the musculature. This is because of conduction delays in efferent and afferent pathways that may last for 100-300 ms (see Miall & King, 2008 for references). Internal models are employed more rapidly than real-world feedback, and provide 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. In motor control, the predictive signal is used to update current representations of the body, as well as to plan and coordinate actions, and help control ongoing movement by issuing corrective commands for deviations from the planned trajectory. Furthermore, the expected sensory feedback (the 'corollary discharge') allows for sensory signals from external stimuli in the environment ('exafference') to be distinguished from sensory signals from the organism's own actions ('reafference').[10]

Therefore, the brain is informed on the current status of the peripheral motor system by integrating sensory information with the predictions of internal models on the outcomes of motor commands. This shows the close coupling of feedforward control with state estimation (e.g. Kawato, 1999; Miall & King, 2008), which the introduction now turns to.

[10] See Webb (2004, box2), for references and discussion of the current use of these terms.

2.2.1.2. State Estimation

In ‘state estimation’, the term ‘state’ refers, in general, to the variables representing the internal condition, the status of a system at a given point in time. For instance, the state of a satellite would include its angular orientation, velocity, and position (Simon, 2006). The problem of computing an optimal estimate of a system’s state from measurements of the system’s inputs and responses was solved by Kalman (1960). The solution is called the ‘Kalman filter’, and has become a standard technique used in engineering for tracking and predicting states of dynamical systems (Kalman & Bucy, 1961; Gelb, 1974). In general, Kalman filters ‘‘continuously test incoming data against expectations’’ (Paulin, 1997, p. 529), i.e. they predict a curve consisting of a set of the most recently sampled values by recurrently applying a correction based on the difference between the most recent value and the prediction for that value. In sensorimotor processing, for instance, the Kalman filter adaptively combines two processes in order to develop a representation of the state of the peripheral motor system: a feedforward process, which uses the efference copies of motor commands along with the current state estimate to predict the next state, by replicating the movement dynamics with a forward model; and a feedback process, which compares the afferent inputs with a prediction of them based on the current state. The sensory error, i.e. the difference between actual and predicted sensory feedback, is used to correct the state estimate generated by the forward model (Wolpert, 1997; Jordan & Wolpert, 2000; Miall & King, 2008).

In order to ensure optimal state estimates, the relative contributions of the feedforward and feedback processes to the final estimate are modulated by the ‘Kalman gain’. This is determined on the basis of the relative variance of the prediction and the sensor signal errors, and defines the extent to which the sensory residual influences the ‘‘a priori estimate’’, i.e. the degree to which sensory input trumps expectation. Grossly put, a Kalman filter processor would be ‘‘top-down’’ to the extent that the Kalman gain is low: the lower it is, the more a final percept is determined by expectation; and it is

‘‘bottom-up’’ to the extent that the Kalman gain is high: the higher it is, the more it is that sensory deliverances override the Kalman filter’s inner dynamic (Grush, 1996, 2004).

2.2.2. Cerebellar Foundations of Feedforward Control and State Estimation

The idea that the CB stores internal models and performs computations of state estimation is a long- and widely-supported one. Space restrictions do not allow for a review of the evidence- the reader is instead prompted to the work cited below.[11] This section briefly presents some highlights illustrating how the CB performs such a role.

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), anatomical, imaging, clinical, and modelling work has established the CB as ‘‘the most likely site for forward models to be stored’’ (Kawato *et al.*, 2003, p. 171). In fact, the idea has been supported for a long time in various forms (Ito, 1970, 1984; Kawato *et al.*, 1987; Paulin, 1989; Miall *et al.*, 1993; Wolpert *et al.*, 1998), to the extent that the CB is nowadays discussed as a neuroanatomical instantiation of computations distinct from those of the cerebral cortex and the basal ganglia (e.g. Doya, 1999; Ito, 2000).

Clinical (e.g. Müller & Dichgans, 1994; Nowak *et al.*, 2007) and imaging studies (e.g. Blakemore *et al.*, 2001; Desmurget *et al.*, 2001; Miall & Jenkinson, 2005) have provided evidence for CB involvement in the prediction of sensory consequences of motor actions, as well as in higher cognitive aspects, such as learning to use a new tool (Imamizu *et al.* 2000).

[11] A book-length review of a range of CB computations of adaptive control is provided in Barlow (2002).

CB dysmetria has offered a characteristic theoretical example of disrupted internal models in the paravermis. In the ‘‘finger-to-nose test’’, failure in which is a typical sign of CB dysmetria (Dow & Moruzzi, 1958), healthy subjects can quickly and accurately bring their index finger to their nose with eyes closed, i.e. without feedback. This presupposes a learning process that converts finger control from a feedback mode to a feedforward mode, as shown in figure 2.1 below. In the initial stage of learning, the motor cortex receives instruction for a finger movement, and acts as a controller of the finger-skeletomuscular system by referring to visual feedback. When the motor cortex sends command signals to the musculoskeletal system of a limb to perform a voluntary movement, an efference copy is sent to the CB model via the cortico-ponto-cerebellar pathway. The predictive output of the CB is relayed back to the motor cortex via the interpositus nucleus. Corticonuclear microcomplexes (CNMCs) in the paravermal cortex-interpositus nucleus develop, through practice, dynamics equivalent to those of the skeletomuscular system. The discrepancies between the predictions of the paravermal CNMCs and the copied skeletomuscular system are transmitted as error signals via the olivo-cerebellar pathway to the CB cortex. CNMCs are thus gradually modified via LTD, becoming a reliable copy model of the system. Ultimately, via sufficient training, the CB internal model replaces the actual skeletomuscular system in receiving commands from the motor cortex, and the external feedback loop may be replaced by an internal CB-based one. In this way, quick, precise, and smooth finger movement is ensured, without reference to sensory feedback (e.g. Ito, 1984, 1993a).

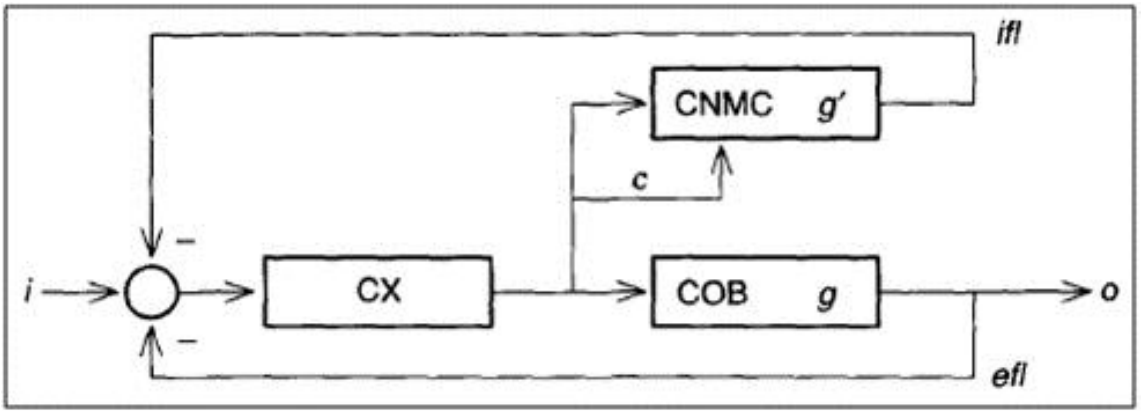


Figure 2.1: The CB as an internal forward model in motor control; c: climbing fiber pathway; CNMC: corticonuclear microcomplex; COB: controlled object; CX: motor cortex; efl: external feedback loop; g, g': motor dynamics; i: input; ifl: internal feedback loop; o: output. Figure reproduced from Ito (1993a, p. 449) with permission © 1993 Elsevier.

In the same vein, the CB has been argued to provide the neuroanatomical analogue of the Kalman filter (Paulin, 1989, 1993, 1997). Evidence has been drawn from CB comparative anatomy, which can not be explained on the basis of a direct CB involvement in motor command generation. The size of the CB and the complexity of the motor apparatus correlate but poorly across species: those with advanced sensory-motor abilities, such as cetaceans, echolocating bats, and electric fishes, in which reafference from motor behavior must be analyzed, have larger CB cortices than their body and brain mass would predict. In such species, then, the CB is associated with purely sensory functions. Primates show a massive expansion of the lateral CB cortex, in parallel with neocortical expansion (see previous chapter, section 1.3.2), but this occurs without an increase in the number of joints or muscles (Paulin, 1993).

The CB has also been shown to filter out expected peripheral feedback as not important to learning. The dorsal cochlear nucleus, a locus much similar to the CB with respect to its architecture and connectivity, has been modelled as a system that nulls reafference: on the basis of the corollary discharge, i.e. the estimated sensory feedback, sensory signals derived from the organism's own actions are stripped off of incoming signals (Nelson & Paulin, 1995). In the same vein, decreased CB activity has been

reported during tickling by oneself, as compared with that shown for the same movements with no tickling involved, and further decreased activity, as compared with externally produced tickling. The explanation put forth has revolved around the same principles (Blakemore *et al.*, 1998).

The poorly coordinated movements in CB ataxia (e.g. Bastian *et al.*, 1996) also demonstrate the significance of state estimation in generating compound motor output. By predicting the future state of a certain muscle, the CB can allow control signals for another muscle to be issued, thus enabling production of simultaneous actions. In the absence of such predictions, the coordination of multiple muscles may only proceed in a seriatim mode, after measurement of the outcome of each command provided by the much slower-arriving external feedback (e.g. Nowak *et al.*, 2002).

Transcranial magnetic stimulation (TMS) experiments have recently provided evidence for the significance of the CB in state estimation in a direct, causal fashion[12] (Miall *et al.*, 2007; Miall & King, 2008). In Miall *et al.* (2007), for instance, participants were asked to make a slow, lateral, untargeted arm movement, before a sudden prompt for them to make a rapid pointing movement towards a static target. While stimulation of other brain areas during the reaching movement did not disrupt performance, TMS over the lateral CB caused reaching errors that were consistent with movements planned on the arm's position about 140 ms previously. CB TMS was thus discussed as having temporarily blocked the CB contribution to state estimation; reaching movements were instead planned on the basis of the residual, out-of-date knowledge of the previous state of the arm.

[12] See sections 4.2 and 4.3 for the capacity of TMS to demonstrate causal relations between task-specific behavior and brain activation.

2.3. Neocerebellar Language 'Dynamics Memory'

The previous section demonstrated that CB circuitry is a well-established foundation of feedforward control and state estimation. I now turn to addressing the gap between psycholinguistics and CB neurocomputational research that this thesis will try to fill.

2.3.1. The Yawning Gap

On the basis of the cytoarchitectural homogeneity of the CB and its reciprocal connectivity with all major subdivisions of the central nervous system (sections 1.2 and 1.3.2), the extension of CB internal models beyond purely sensorimotor control was foreseen quite early (e.g. Ito, 1993a, b; Paulin, 1993; Kawato, 1999). The CB has thus been proposed to copy the dynamics of any kind of neural control exerted by the central nervous system, not only upon bodily (e.g. limbs) or external (e.g. tools) manipulanda, but also upon brain structures per se. In the case of language and thought processes, this is putatively achieved in the following way: frontal cortical areas (in the case of language, Broca's area) issue control signals to the temporo-parietal areas (in the case of language, Wernicke's area); a CNMC acquires the dynamics of this cortico-cortical processing, and becomes a reliable copy of particular operations on the representations of the temporo-parietal cortex (figure 2.2). In this way, the frontal cortical areas, i.e. the controller, may operate on the CNMC instead, thus avoiding the costlier exploration of the temporo-parietal areas (e.g. Ito, 1990, 2000, 2008; see sections 2.3.4.2 and 2.4.4.6 here).

Despite the plausibility of the idea, research in these domains remains stagnant. This is because of the poor insight of CB neuroscientists into the particular language processes which CB circuitry may copy the dynamics of – the former comprise the ontology of psycholinguistics. The yawning gap between psycholinguistics and CB neurocomputations is made apparent in figure 2.2: terms like ‘‘thought model’’, ‘‘words’’, ‘‘grammars’’, ‘‘action’’, ‘‘speech’’, indicate the very vague and speculative character that such reflections are constrained to take in the absence of interdisciplinary discourse. In the same vein, studies of CB-induced agrammatism (see section 1.3.3 here) are cited as suggestive evidence for the participation of NCB internal models in language processing (e.g. Ito, 2000, 2008). However, such cases may at most support the idea of NCB involvement in language processing in a general fashion. If CB lesions result in the same language deficits that cortical lesions do, then these language deficits cannot provide direct evidence for NCB involvement in language processes particularly via internal models. CB-induced agrammatism has other explanations that are equally, if not more, convincing (cerebro-cerebellar diaschisis, verbal working memory impairment; section 1.3.3); in fact, it is the exception, rather than the rule in CB pathology (Daum *et al.*, 2001). Proof for the involvement of internal models in language processing thus cannot be drawn from the cortical-like deficits that CB impairments may induce. On the contrary, very particular CB language deficits have been observed in the absence of such cortical-like signs (e.g. Fiez *et al.*, 1992; Gebhart *et al.*, 2002), which deserve closer investigation (section 2.4). However, even these cases have been described on the basis of motor-based metaphors for NCB higher cognitive function, with terms as abstract as those of ‘‘mental movement’’, or of attaching a verb to a noun-context (Thach, 1996a; see Gebhart *et al.*, 2002).

With respect to such demonstranda, then, the recent introduction of internal models in psycholinguistic research stands as a very fortunate contingency. In the remainder of this section, I will show that independent work in cognitive science and psycholinguistics can help identify an ontology of linguistic mechanisms, and replace abstract speculation with neurocomputationally-based CB models and experimental hypotheses on language processing.

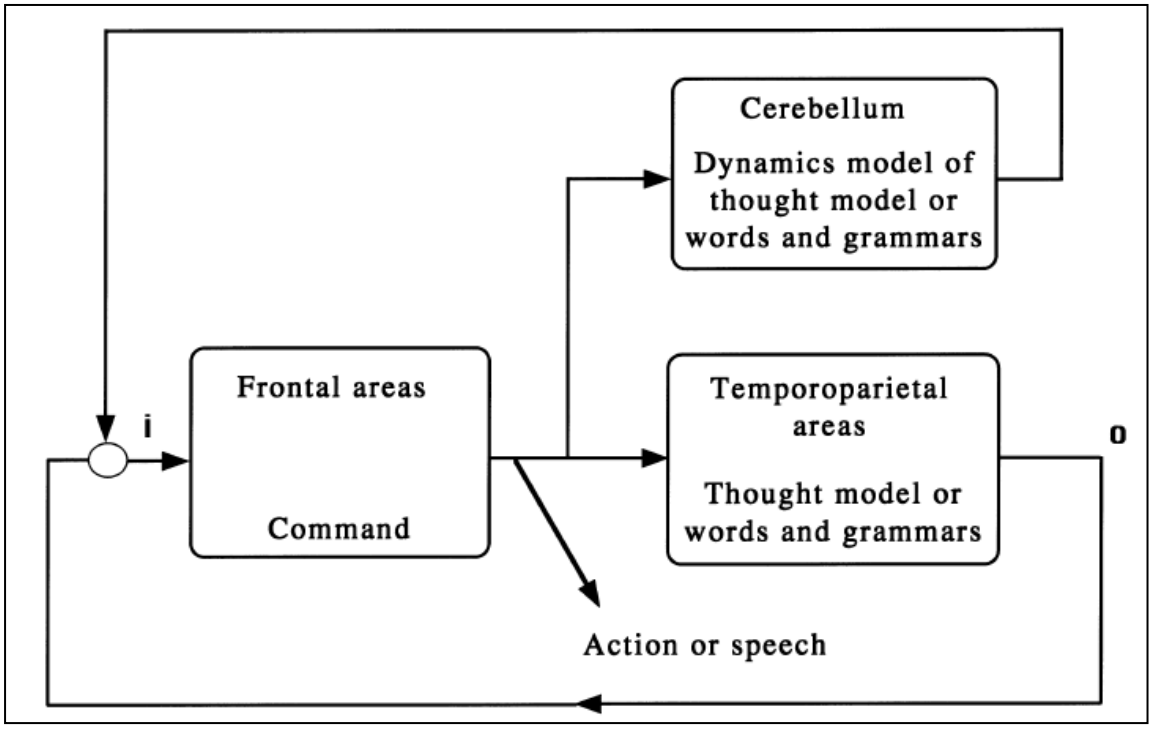


Figure 2.2: Involvement of a NCB internal model in language and thought processing; i: input; o: output. In a feedforward fashion, the NCB becomes a reliable copy of the controlled object (temporo-parietal areas). Frontal areas (the controller) may now act on the NCB instead, in a faster and less attended fashion. Figure taken from Ito (2000, p. 158) with permission © 2000 MIT Press.

2.3.2. Feedforward Control and State Estimation in Language Processing

Interestingly, work on internal models has been accommodated in modern cognitive science and, in particular, psycholinguistics, outwith the context of CB neuroscience. The first significant step in that direction was taken within the scope of ‘‘the emulation theory of representation’’ (Grush, 1996, 2004). ‘‘Emulators’’ are neural circuits that operate on the same principles as feedforward models and Kalman filters. They thus act as models of the body and the environment, and, during sensorimotor behavior, they provide expectations for sensory feedback in a pre-emptive and enhancing fashion. They can also be driven off-line on the basis of efference copies, in order to produce imagery.

These emulative computations have been accommodated in the literature on the action-based representations employed in perceiving conspecifics. In a covert, imitative fashion, perceivers are held to use their own apparatus for the production of the same actions observed. This apparatus crucially involves emulators providing predictions for the sensory consequences of these actions. These predictions are adaptively involved in the perceptual process, maximizing, in general, its speed and quality (see Wilson & Knoblich, 2005, for details).[13]

In the line of such work, Pickering and Garrod (2007) propose that equivalent production mechanisms are employed in language perception-comprehension in a facilitatory manner. Production, in other words, holds a causal role in efficiently processing noisy and/or ambiguous linguistic input and in generating predictions:

[13] Characteristically, in the work of Grush (1996, 2004), the CB retained its status as the most plausible candidate for the instantiation of such computations. Subsequently, though (e.g. Wilson & Knoblich, 2005), attention has been shifted to mirror neurons, with no reference to the CB (see section 2.5 for further discussion).

> […] comprehenders use prediction and imitation to construct an 'emulator', using the production system, and combine predictions with the input dynamically (*ibid*, p. 105).

Their processor thus uses a ''production-based language emulator'', which is ''controlled by feedback from a Kalman filter'' (*ibid*, p. 108), weighing predictions against analysis of the input at each step at all linguistic levels (phonology/syntax/semantics). In cases where the prediction is strong and the input noisy (''low Kalman gain''), the emulator exerts strong top-down influence to the ''input analysis system''. In that way, noisy percepts may be enhanced based on the expected input. Furthermore, ''comprehenders can get ahead of themselves and have more time to keep up with what they are encountering'' (*ibid*, p. 105).

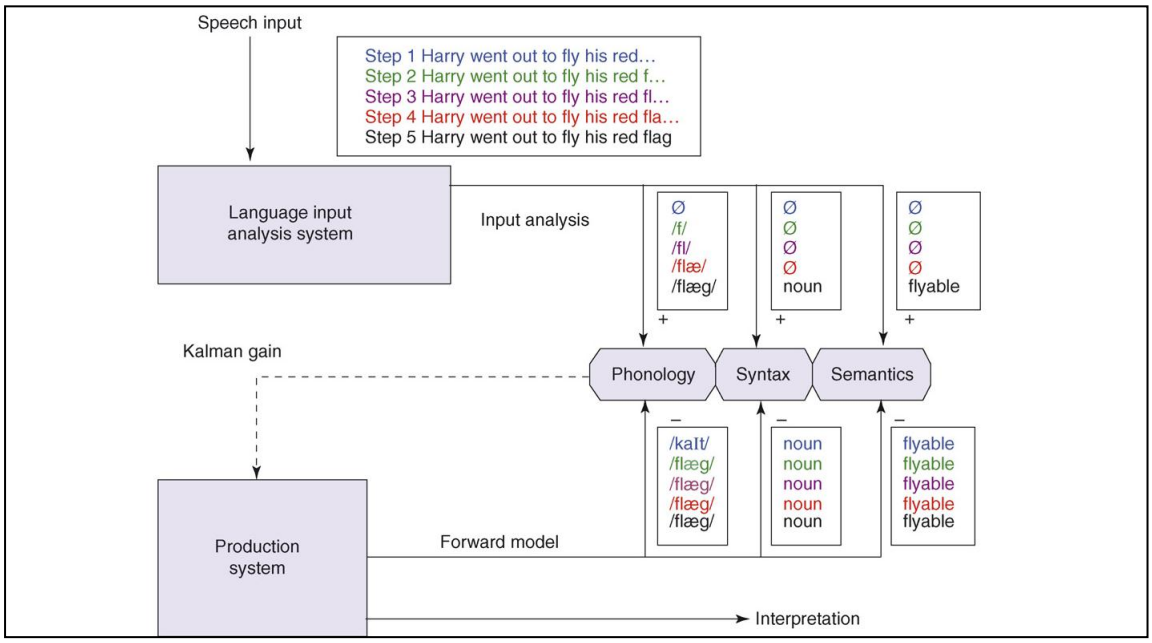


Figure 2.3: State estimation and feedforward control in language comprehension in Pickering and Garrod's (2007) processor. The emulator is controlled by feedback from a Kalman filter, which weighs predictions against analysis of the input at each step. Five steps are illustrated in comprehending the end of the sentence 'Harry went out to fly his red flag.' The outputs of the input analysis system and those of the internal model are shown in the same color for three different levels of prediction (phonology, syntax and semantics) at each step. Figure reproduced from Pickering & Garrod (2007, p. 108) with permission © 2007 Elsevier.

2.3.3. The Model

As suggested above, the introduction of state estimation and feedforward control in language processing provides the keystone for the articulation of concrete NCB neurolinguistic models.[14] Above all, it may help give a first answer to the question of how NCB computations may contribute to language processing, along with a set of falsifiable experimental hypotheses. The Neocerebellar Kalman Filter Linguistic Processor (henceforth NCBKFLP; figure 2.4 below), as first presented in Argyropoulos (2009), is explicitly formulated on the basis of the Pickering and Garrod (2007) processor and on the functional principles of CNMCs (Ito, 1984; section 1.2.4 here). It thus offers a proxy for synthesizing insights from psycholinguistic processes and CB state estimation-feedforward control, and provides a grounding for the considerations to follow. [15]

Linguistic internal models of the NCBKFLP receive a copy of the input processed at each stage by the cortically instantiated[16] ‘‘input analysis system’’ (Pickering & Garrod, 2007) along with the operations performed on it, and output a prediction of the next stage that the system will enter. CB predictions would be transmitted by the ventrolateral neodentate nucleus via the phylogenetically newer parvocellular part of the red nucleus and the ventrolateral nucleus of the thalamus to Broca’s area. The discrepancies between the actual outputs of the input analysis system and the models’ predictions provide the error signals training the latter. These would be conveyed by climbing fiber signals, inducing LTD at the synapses of Purkinje cells-parallel fibers of the NCB cortex (section 1.2.4), thus training NCB predictions via trial and error. In low Kalman gain contexts, where input analysis is conducted in noisy conditions and/or strong NCB predictions are transmitted, NCB output may override

[14] Feedforward control has been used in the past in psycholinguistic models, but with much narrower application, e.g. in auditory feedback for speech production (Perkell *et al*., 1997).

[15] Covert imitative language production in perception is compatible with, but not fundamental for the model. See section 2.5 for further discussion.

[16] See section 2.3.4 below for the cortical instantiation of the input analysis system.

and/or bypass the output of cortico-cortical processing.[17] The strength of the transmitted prediction may be at least partly determined by the strength of inhibition that the neodentate (constantly excited by collaterals of mossy and climbing fibers; section 1.2.3) receives from the Purkinje cells of the NCB cortex. The strength of such inhibition is a function of the 'punishment' that Purkinje cells receive via LTD for their participation in a suboptimal prediction (section 1.2.4).

However, the model assumes that the inferior olive receives projections both from CB nuclei and from the temporo-parietal cortex. Despite the evidence for an inhibitory projection from CB nuclei to the inferior olive (see Ito, 2008, for references) there is no evidence for a direct projection from the temporo-parietal cortex. In fact, the inferior olive receives little direct input from the cerebral cortex, and its major input arises from the red nucleus that conveys mostly sensorimotor information. However, the inferior olive also receives associative cortical input indirectly from brainstem reticular nuclei and through the zona incerta. The latter provides input from the rostral cingulate cortex (BA 24), prefrontal cortex (BA 9/46d), the posterior parietal cortex, and the medial prestriate cortex (Shah *et al.*, 1997). Thus, the associative projections to the zona incerta, which in turn projects to the inferior olivary nucleus, maintain the possibility that CB microcircuitry may perform the same computations on higher functions (Schmahmann, 2000, 2001). In that way, then, the predictions of the NCB linguistic internal model can be compared with the actual output of the cortical input analysis system. As described above, the discrepancies between the two may train the model to adequately copy the dynamics of instances of input analysis. Section 2.3.4 below describes in more detail what properties of input analysis may be copied, and how they may be overridden and/or bypassed by NCB internal models.

[17] No particular commitment is made with respect to the temporal arrangement of these two outputs (see section 2.4.4.1 for discussion).

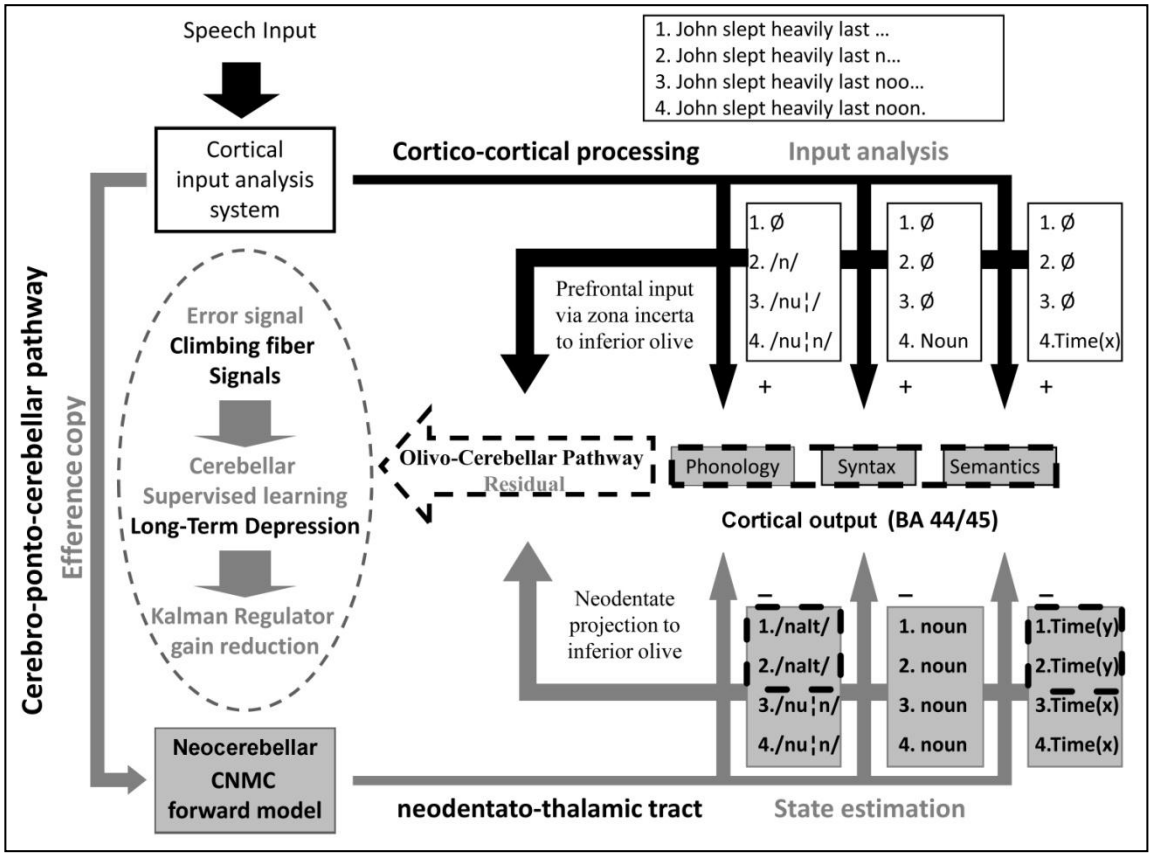


Figure 2.4: The NCBKFLP in perception-comprehension of the end of the sentence 'John slept heavily last noon'. The predictions of a NCB internal model are dynamically combined with the outputs of the cortical input analysis system. For the sake of simplicity, only one prediction is presented at each stage. Mossy fiber inputs from the pontine nuclei convey, via parallel fibers, information on the part of the sentence processed by the cortical input analysis system to Purkinje cells in the NCB cortex and the ventrolateral dentate nucleus. The latter, via the parvocellular red nucleus and the ventrolateral thalamus, signals the expected state of input analysis back to its language-related cortical outputs. Because of the strong semantic association between 'sleep' and 'night', the CB internal model outputs, in steps 1 and 2, strong predictions for the phonological, morphosyntactic and semantic representations of 'night', which are, however, not verified by the input analysis system. The discrepancies between the two are conveyed as error signals back to the NCB cortex, via the climbing fibers, which receive indirect connections from the cortical language-related areas via the zona incerta and direct projections from the dentate nucleus. This residual is used to train the internal model, via LTD at the parallel fiber-to-Purkinje cell synapses. Figure adapted from Argyropoulos (2009, p. 194). [18]

[18] It should be clarified for the discussion to follow that Kalman filters per se do not reduce the Kalman gain. Instead, a 'Kalman regulator' provides the adaptive mechanism whereby the Kalman gain and the randomness of the residual can be used as cues to improve performance (M. Paulin, personal communication, 25 October, 2007). Cerebellar LTD provides the basis for such mechanisms (sections 1.2.3.3 and 1.2.4).

2.3.4. Neocerebellar Linguistic Associative Memory

Expectedly, the model has so far not concentrated on any particular details on language processing in cerebral loci, i.e. what was called the 'cortical input analysis system', or the 'cortico-cortical processing' modality (figure 2.4). However, the very nature of the computations of the NCB, i.e. copying the dynamics of processes conducted elsewhere in the brain, makes it almost instructive that cortical linguistic operations be considered. In this fashion, briefly discussing how cortico-cortical language processing may be performed would provide insight into the nature of the control signals that Broca's area would exert on the NCB microcomplexes.

Some basic insight into these questions may be offered by the ''Memory, Unification, and Control'' model (henceforth, MUC; Hagoort, 2003). In fact, incorporating the CB and the basal ganglia[19] in such a model has been identified as an outstanding desideratum (Hagoort, 2005, p. 422). A brief description of the MUC is now in order.

2.3.4.1. The 'Memory, Unification, and Control' Model

The MUC grounds language processing on the basis of the interaction of three neurocognitive components: the 'Memory' component, distributed over the left temporal cortex, pertains to multi-level language information stored in long-term memory (phonological/phonetic features and semantic features of words, as well as syntactic features and frames), as well as to the corresponding retrieval operations. The 'Unification' component performs the selection and integration of information retrieved from the Memory component into a representation of larger structures, for all phonetics/ phonology, syntax and semantics. Broca's area and the adjacent cortex (left inferior

[19] See section 3.7 for an argument on basal ganglionic involvement in language processing in the light of grammaticalization operations.

frontal gyrus) are assumed to instantiate the Unification component. Finally, the 'Control' component, instantiated in the anterior cingulate cortex and dorsolateral prefrontal cortex (BA 46/9), is assumed to relate language to the extra-linguistic communicative settings, by involving attentional control in language interaction. This is the case in temporospatial deixis (e.g. 'here', 'now'), conversational turn-taking, or language-selection in the case of bilingualism (Hagoort, 2003).

2.3.4.2. Neocerebellar Feedforward Control in the MUC

The idea of unification operations performed by the left inferior frontal gyrus on items stored in and retrieved by the left superior temporal cortex (Hagoort, 2003) provides ideal grounds to explore the processes of the MUC in terms of recent control-theoretic accounts – i.e. in terms of 'control signals' of the controller-Broca's area being sent to the controlled object-Wernicke's area (e.g. Ito, 2000, 2008; also here, section 2.3.1). In fact, this conceptualization of Broca's area as the controller and Wernicke's area as the controlled object has been complemented by recent considerations on the neuroanatomical correlates of the ''instructor''. The instructor, i.e. the part of a control system that gives a goal towards which a control system should work, is putatively instantiated in loci including the anterior cingulate gyrus. In thought contexts, for instance, it has been proposed to supply an instruction to the controller in the prefrontal cortex, which in turn sends control signals to a mental model stored in the temporo-parietal cortex (Ito, 2008).

There is thus a striking parallelism between the MUC components (figure 2.5.A) and those independently identified in Ito's (2008) latest reflections on feedforward control of higher cognitive operations (figure 2.5.B). The anatomical co-occurrence of those two tripartite distinctions makes their integration quite tempting: the Control-instructor component, the Memory-controlled object component, and the Unification-controller component form a system that may now invite NCB computations (figure 2.5.C). Upon the formation of an adequate forward model in the NCB, Broca's area may now perform its unification operations on the model, and not on Wernicke's area.

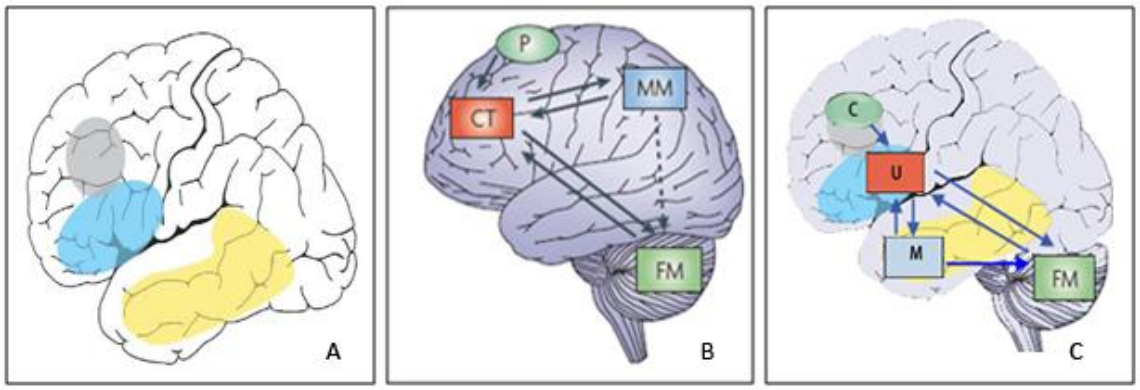


Figure 2.5: A synthesis of the MUC framework with NCB internal models. A. The three components of the MUC model. Memory (yellow) in left temporal cortex, Unification (blue) in LIFG, and Control (grey) in dorsolateral prefrontal cortex. Figure adapted from Hagoort (2005, p. 421) with permission © 2005 Elsevier. B. Feedforward control of thought processes. P: Instructor. CT: Controller. MM: Mental Model. FM: Forward Model. The CB learns the dynamics of the control signals sent to the mental model by the prefrontal cortex. The same principles apply to the control of Wernicke's area by Broca's area (Ito, 2000, 2008). Figure reproduced from Ito (2008, p. 306) with permission © 2008 Nature. C. Emulation of the Memory component by the CB. C: Control. U: Unification. M: Memory. FM: Forward Model. In a cortico-cortical processing modality, upon instruction by the Control component, the Unification component operates on the linguistic units of the Memory component. The CB forward model receives a copy of the control signals of the Unification component and feeds back to it its prediction for the outcome of the cortico-cortical operation. Future predictions are subsequently trained on the basis of the discrepancy between the actual outcome of the cortical operation and the CB prediction. In re-occurring unification operations, the CB forward model provides reliable predictions, and Unification operations are thus performed on the CB look-up table, and not on the Memory component.

2.3.4.3. The Categorical-Associative Trade-off

In what way would the unification operations performed by Broca's on Wernicke's area differ from those conducted by Broca's area on the NCB? And how could the adaptive significance of the latter be grounded in psycholinguistic terms? In answering this question, it is important to first consider the two fundamental ways in which CB emulation of cortical processes differs from the very algorithms the latter employ.

CB computations differ from cortical ones in ''speed'' and ''flexibility'' (Doya, 1999; Ramnani, 2006). On the one hand, CB mechanisms show greater speed, accuracy, and automaticity as compared with the emulated cerebral ones, which rely on feedback control. A striking piece of evidence is the fact that the CB responds to sensory input even before the very primary sensory cortical areas (see Ramnani, 2006, for references). On the other hand, CB learning algorithms show poor capacities to generalize across situations, and are limited to particular contexts (*ibid*). In the motor domain, for instance, it has been independently suggested that the CB would acquire the coordination of only specific muscle activities. Similarly, in 'eye-blink conditioning' (see below), redirection of the conditioned stimulus from the eye to a finger would not involve transfer of the conditioned response in the form of flexion or extension of the finger (Hazeltine & Ivry, 2002). This is why the CB ''may not be designed for the development of abstract representations'' (*ibid*, p. 1980).[20] The same argument is also found in the construal of the CB as an instantiation of supervised learning algorithms, where learning proceeds on the basis of pre-specified input-output pairings (Doya, 1999; see section 2.2.1.1 here).

[20] However, there is no reason to commit CB inputs and outputs to a particular level of abstractness. The CB would certainly not be able to generalize over its inputs, or to retrieve particular inputs as instances of more abstract representations. However, it may still receive abstract categories as inputs from the cerebral loci after the appropriate cerebral retrieval operations, and may transmit such representations back to cortical areas. Suffice it for these representations to be reliable predictors of particular outputs of a copied cortico-cortical process (Ohyama *et al*., 2003). Associations between abstract categories, I will argue below, might involve NCB circuitry, as in the case of the N-V-N pseudosyntactic heuristic, or of schematic transmission (section 2.4.4.5 below).

On the contrary, cerebral algorithms are assumed to abstract rules from context in an unsupervised learning fashion, and flexibly apply them to different contexts (*ibid*). In unsupervised learning, the algorithm is presented with examples from the input data set, and is called to fit these observations in a model – with no pre-specified output, much unlike supervised learning. In clustering, for example, a characteristic method of unsupervised learning, the algorithm assigns a set of observations into subsets (‘clusters’), so that observations in the same cluster are similar according to some criterion (e.g. Pothos & Chater, 2002). However, CB computations are advantageous in constrained behavioral contexts: if the same rule is applied to the same context in a deterministic fashion elsewhere in the brain, then CB algorithms may arbitrarily replicate these reoccurring input-output mappings (Doya, 1999). In this fashion, the CB may provide faster solutions with fewer resources, like a ‘‘shortcut’’ or a ‘‘look-up table’’ (*ibid*; Ramnani, 2006; section 2.4.4.2 here).

In what follows, I argue that this ‘speed/flexibility’ trade-off between CB and cerebral computations is reflected in language processing with the ‘categorical/associative’ distinction, respectively. In particular, I would like to propose that NCB internal models store multi-level linguistic associative relations, and that the Kalman filter of the NCBKFLP regulates the trade-off between the cortico-cortical, categorical modality, and the associative, cortico-cerebellar modality. In contexts of increased predictability (‘low Kalman gain’), the associative, cortico-cerebellar computations are maximally involved, overriding and/or bypassing the cortico-cortical ones.

Evidence for the fundamental role of the CB in the formation and storage of associative memory traces is indeed found in a wide range of paradigms, from classical conditioning to cognitive tasks. Classical conditioning, in the form of the ‘nictitating membrane response’, also known as ‘Pavlovian eyelid conditioning’, represents the most basic form of long-term associative memory formation. In such tasks, subjects are trained to generate an eye-blink (conditioned response) to an initially neutral stimulus,

e.g. a warning tone (conditioned stimulus), which is temporally paired to and thus predicts another stimulus, e.g. a corneal air puff (unconditioned stimulus). Before conditioning, the tone alone does not elicit an eye-blink, but the unconditioned stimulus does. During conditioning, however, the tone becomes a reliable predictor of the unconditioned stimulus, so that, after conditioning, the tone alone evokes an eye-blink. As training proceeds, there is a gradual increase in both the probability and the amplitude of conditioned responses (Hazeltine & Ivry, 2002).

If the CB operates as a feedforward control system, then it should display associative learning, whereby output is adjusted for sensory inputs that reliably predict errors (Ohyama *et al.*, 2003). Indeed, the CB has long been established as the fundamental site of classical conditioning (e.g. Thompson *et al.*, 1997; Christian & Thompson, 2005) in both comparative (e.g. McCormick & Thompson, 1984) and human clinical studies (e.g. Daum *et al.*, 1993; Timmann *et al.*, 1996; Woodruff-Pak *et al.*, 1996). Deficient eye-blink conditioning has also been reported in populations involving CB pathology, e.g. schizophrenics (e.g. Brown *et al.*, 2005), or dyslexics (Nicolson *et al.*, 2002). In particular, the conditioned and unconditioned stimulus representations are conveyed to the CB via the mossy and climbing fibers, respectively, and CB output via the deep nuclei expresses the conditioned response. At first, the anterior interpositus nucleus is inhibited, because the conditioned stimulus is passed to the CB cortex via the mossy fibers, which activate the inhibitory Purkinje cells. However, when specific mossy fibers inputs (conditioned stimulus) reliably co-occur with an error signalled by a climbing fiber input (unconditioned stimulus), with a specific time delay, LTD punishes those synapses that inhibit the anterior interpositus, and CB output is adjusted to occur at the appropriate time (e.g. Ohyama *et al.*, 2003). Such findings have been extended with reports of CB impairments in learning associations between visual cues and motor commands (Canavan *et al.*, 1994; Tucker *et al.*, 1996), and MRI-assessed CB volume has been found to correlate with associative learning (Woodruff-Pak *et al.*, 2001).

More recently, CB involvement has been shown in non-motor associative learning. In one such study, patients with isolated CB degenerative disease were asked to learn the association between six pairs of colors and numerals by trial and error. In comparison with normal control subjects, it took the patients significantly longer to learn the correct associations between colors and numerals, and they were impaired in recognizing them later on, while two patients showed no associative learning effect at all. Additionally, the patients were also tested in two control tasks in order to control for the influence of their motor performance in the cognitive task: neither of those tasks correlated substantially with their poor cognitive performance (Drepper *et al.*, 1999).

If anything, then, the involvement of CB structures in such multi-level associative computations of temporally paired events strongly supports the proposal here – i.e. that in language processing, temporally paired relations between linguistic events can be acquired, stored, and adapted in NCB circuitry. In other words,

> [w]hether activated by a tone, the position of a limb or some cognitive process, the cerebellum will process mossy fiber inputs as arbitrary signals that do or do not reliably predict events (Ohyama *et al.*, 2003, p. 226).

The proposal here is thus that CB ‘‘dynamics memory’’ (Ito, 1993a, b) is involved in the language domain with the formation and storage of associative linguistic relations. At an elementary level, the capacity of the NCB to prepare the language-related cortical loci for upcoming events should manifest itself in lexical ‘‘associative’’, and not ‘‘semantic’’ priming (see also section 2.6.2 below):

> [a]ssociative relatedness is a normative description of the probability that one word will call to mind a second word […] Associative relations are assumed to reflect *word use* rather than word meaning (e.g. ‘‘needle-thread,’’ ‘‘spider-web’’). The source of these associations might be

> *temporal contiguity* in verbal or written language […] or *co-occurrence* within propositions […] On the other hand, semantic relatedness reflects the *similarity* in meaning or the *overlap in featural descriptions* of two words (e.g. ‘‘whale-dolphin,’’ ‘‘duck-chicken’’) (Thompson-Schill *et al.*, 1998, p. 440; *italics* mine).

However, this widely used ‘associative/semantic’ divide silently confounds the form/meaning distinction with that of the paradigmatic/syntagmatic axis, or, more appropriately, of categorical/associative relations. The ‘form/meaning’ distinction is made between different levels of linguistic analysis, while the ‘categorical/associative’ one describes two different relations linguistic items might be involved in. In other words, meaning relations are not necessarily categorical in nature, and form relations between words are not necessarily associative. Such a confound would exclude the co-occurrence relations between semantic representations of (often co-occurring) words, such as thematic-role (θ-role) -bearing verbs and their corresponding θ-role-filling NPs (e.g. ‘customer-buying’, ‘table-eating’) on the one hand, and word-lexeme-lemma categorical relations (/running/-RUN) on the other.[21] In an attempt to refine the terminology involved and expand the corresponding ontology, a 2 x 2 classification of relations is summarized instead in table 2.1. Associative relations may thus hold in both formal and semantic levels. In pair-wise lexical relations, ‘formal-associatively related’ items would most aptly correspond to Hutchison’s (2003) ‘‘forward phrasal associates’’ (e.g. ‘pigeon-hole’, ‘couch-potato’). Apart from θ-role-based relations, ‘semantic-associative’ ones would also involve cases like the ‘‘script relation (orchard-apple)’’, or the ‘‘instrument (broom-floor)’’ relations in Hutchison (2003), or the ‘‘functionally related-associated instruments’’, e.g. ‘‘pram-baby’’ and ‘‘bandage-wound’’ in Moss *et al.* (1995). This ontological distinction will be necessary in formulating the experimental predictions in the TMS experiments reported here (chapters 5-7).

[21] Unavoidably, processing lexical items that are related in a ‘formal-categorical’ fashion is confounded with repetition and semantic-categorical priming. This is because lexical items in the same formal network involve the same morphemes.

Direct evidence for NCB involvement in formal-associative computations is in fact demonstrated in the first TMS study conducted here (chapter 5). For semantic-associative relations, supportive evidence is provided by the performance of NCB patients in verb-generation tasks (section 2.4.4.5), as well as by the findings of the third TMS study reported here (chapter 7). On the other hand, categorical representations will be shown to remain intact in CB patients, both in the semantic and in the formal syntactic level of processing (see section 2.4.4.4).

Level	**Relation**	
	Associative	**Categorical**
Formal	**Formal-Associative** 'gift-horse' 'storm-teacup'	**Formal-Categorical** 'running'- 'runs' 'run'- 'ran'
Semantic	**Semantic-Associative** 'customer-buying' 'chef-cooking'	**Semantic-Categorical** 'fruit-apple' 'apple-pear'

Table 2.1: The associative-categorical, formal-semantic quadripartite distinction.

At this point, I would like to introduce ''constructions'' as the control objects for the unification-control signals of Broca's area. This would be in equivalence with mental models in the temporo-parietal areas being controlled by the prefrontal cortical controller in thought processes (Ito, 2000, 2008).[22] Constructions are traditionally defined as ''stored pairings of form and function, including morphemes, words, idioms, partially lexically filled and fully general linguistic patterns'' (Goldberg, 2003, p. 219; see table 2.2 for examples). With respect to the representation of constructions, then, the Memory component would be seen as instantiating what Goldberg (2003) calls a

[22] Constructions will also be a valuable tool in exploring the neurolinguistic foundations of grammaticalization processes, since constructions are at the same time the units on which such changes operate (chapter 3).

‘‘construct-i-con’’, in the sense that ‘‘[t]he totality of our knowledge of language is captured by a network of constructions’’ (*ibid*, p. 219). In fact, their representation reflects basic properties of the way semantic nodes of lexical items are organized into networks. For instance, similar referents and similar constructions are represented by nodes close to one another, and processing of one such node involves spreading activation to adjacent nodes (Goldberg, 2003; Diessel, 2004). More abstract, schematic constructions, also called ‘‘constructional schemas’’ (e.g. Bybee, 1995) stand higher in the constructional hierarchy and transmit, via mechanisms of default inheritance, their syntactic and semantic specifications to less schematic ones. Less productive-canonical patterns are captured at different midpoints of these hierarchies (e.g. the idioms in table 2.2). Accordingly, language-specific generalization relies on the induction of more abstract schemata from more concrete constructions.

Construction	Form/Example	Function
Morpheme	e.g. *anti-, pre-, -ing*	
Word	e.g. *Avocado, anaconda, and*	
Complex word	e.g. *Daredevil, shoo-in*	
Idiom (filled)	e.g. *Going great guns*	
Idiom (partially filled)	e.g. *Jog* (someone's) *memory*	
Covariational-Conditional construction	Form: The Xer the Yer (e.g. *The more you think about it, the less you understand*)	Meaning: linked independent and dependent variables
Ditransitive (double-object) construction	Form: Subj [V Obj1 Obj2] (e.g. *He gave her a Coke; He baked her a muffin*)	Meaning: transfer (intended or actual)
Passive	Form: Subj aux VPpp (PP_{by}) (e.g. *The armadillo was hit* by a car)	Discourse function: to make undergoer topical and/or actor non-topical

Table 2.2: Examples of constructions of different sizes and complexity; function is provided wherever not transparent. Table adapted from Goldberg (2003, p. 220) with permission © 2003 Elsevier.

The ‘flexible’, categorical manner in which cortical learning may proceed would be reflected in language processing by the notion of ‘‘schematization’’ (Langacker, 2000). This involves the analogical construction of schemata by extraction of common features from the language input data. The particular features extracted from distributional data reinforce each other and give rise to more abstract representations

(Diessel, 2004). Capacities to generalize over data were seen above to rely selectively on cortical unsupervised learning algorithms, and not on CB supervised ones. The linguistic nature of the learning input would place the relevant computations in the temporo-parietal cortical Memory component (see section 2.4.4.4 for evidence).

In sentence processing, then, strings are processed and recognized as potential members-instantiations of different constructions. Via vertical, categorical links, the corresponding schematic constructions are activated in the cerebral-Memory component, along with their syntactic-semantic specifications; these provide the constraints according to which unification with subsequent input strings proceeds. On the contrary, the CB is blind to such vertical-inheritance relations. Instead, it stores associations, i.e. instances of the dynamics of particular unification operations. Over time, it becomes a reliable look-up table storing co-occurrences between particular items, or between particular items and their structural analyses. In that way, it provides a constrained search space within which the controller, i.e. Broca's area, may immediately retrieve unification candidates. This can be done without consulting the Memory component, i.e. without checking the membership of a particular construction in a particular schema. For now, a brief example could be given with the ditransitive construction of the sentence 'He gave her a coke' in table 2.2 above. The cortical input analysis system may assign a structural representation [Obj] to the pronoun 'her' only after the activation of the schematic construction [GIVE Obj_1 Obj_2] and/or the even more schematic [$V_{(ditransitive)}$ Obj_1 Obj_2]'. On the contrary, the frequent co-occurrence of the phonological representations of 'give' and 'her' in discourse, as well as the strong association between the phonological input 'gave her' and the structural output [$gave_{(V)}$ $her_{(Obj)}$] may invite maximized involvement of NCB 'dynamics memory'. The latter will bias, override, and/or even bypass cortico-cortical processing, assigning this structural representation without and/or before activation of the cortically-stored schemata above.

In this section, the argument was made predominantly with reference to computational properties of the different neuroanatomical components involved, and with minimal psycholinguistic insight. In the discussion that follows, those associative, non-categorical, possibly pre-emptive multi-level linguistic computations that the CB instantiates will be clearly illustrated in the light of evidence in the literature and in the context of particular psycholinguistic models.

2.4. Support for the Model

Support for NCB emulation in language processing predominantly relies on the anatomical and computational properties of the CB, and only secondarily on empirical evidence. This is owing to both the 'cerebral chauvinism' in the interpretation of NCB-induced cognitive deficits, as well as to the early infancy of computational approaches to NCB cognitive functions (sections 1.3.2, 1.3.3, and 2.3.1). However, certain aspects of NCB linguistic functions have been noted in the clinical and imaging fields to reflect NCB computations, yet in a sporadic and speculative way. It is thus instructive to examine such evidence from the NCBKFLP perspective.

2.4.1. Error Detection

Evidence for NCB error signalling and error-driven learning in the linguistic domain was first provided in Fiez *et al.* (1992). Their NCB patient showed deficient supervised learning in a concurrent discrimination word learning task,[23] and likewise failed to detect instances of their erroneous performance in word-generation tasks. In a PET study of sentence processing (Stowe *et al.*, 2004), significant NCB activation was found at the point of reanalysis in reading ambiguous as opposed to unambiguous sentences. However, this finding was interpreted within a framework very limited with respect to the overall computations of the CB, whereby the CB is primarily involved in error detection and signalling (Jueptner & Weiller, 1998). On the contrary, the computational work on the CB and the NCBKFLP hypothesis encompass error detection and signalling as only one of CB functions. As shown in figure 2.4 above, the discrepancies between the predicted, neodentate-based, and the actual, cortical outputs

[23] In each of the 20 pairs of words used, one word was assigned to be the 'positive member'. Subjects learned the positive word in each pair through trial and error: after indicating their selection by pointing, subjects were informed whether they were correct. Word pairs were repeated until a subject made no errors on four consecutive blocks, up to a maximum of twenty (Fiez *et al.*, 1992).

of a linguistic operation are transmitted via the olivo-cerebellar pathway, training, via LTD, the CB cortex.

2.4.2. Practice-Dependent Optimization

If the CB instantiation of the Kalman filter is disrupted, then the Kalman gain cannot be lowered on the basis of the residuals, and the CB model may not become a reliable copy of the target process. Lack of practice-induced facilitation is characteristic in CB patients (see section 3.5.3), and has also been shown in linguistic tasks, yet with virtually no reference to the internal models framework. In such tasks, patients reduce their response latencies poorly across blocks in comparison with normal controls. For example, in a verb-generation task, the reaction times of the patient in Fiez *et al.* (1992) did not improve as a function of practice, whereas normal control subjects reduced theirs by approximately 30%. Similar findings have been observed in more recent studies on NCB patients (e.g. Gebhart *et al.*, 2002). Recently, CB transcranial direct current stimulation (tDCS) has been shown to specifically impair the practice-induced proficiency increase in a verbal working memory task. On the contrary, stimulation over the prefrontal cortex induced an immediate change in task performance, sparing, though, the practice-related proficiency changes (Ferrucci *et al.*, 2008).

2.4.3. Noise-Resistant Speech Perception

Since state estimators function as optimal noise-rejecting filters for signals modelled as outputs of dynamical systems, the CB might in certain circumstances induce perceptual enhancements. For example, CB circuitry could be employed for a mechanic's 'tuned ear' to detect acoustic signatures of particular mechanical breakdowns in an engine (Paulin, 1997, p. 527). This should also hold for the outputs of

the NCBKFLP to the language-related cortical loci. Some first suggestive evidence for such involvement in language percepts can be drawn from dyslexic populations, where CB pathology is arguably involved (e.g. Nicolson *et al.*, 2001; see section 3.5.3 here). The pronounced difficulties of dyslexics in perceiving speech in noisy contexts (e.g. Sperling *et al.*, 2005; Ziegler *et al.*, 2009) have been speculatively attributed to deficient CB internal models in speech processing (Ito, 2008). [24]

2.4.4. Sentence Processing Habits

So far, the discussion topics have been defined by some general aspects of CB computations (practice-induced optimization, error signalling, perceptual enhancement) that are reflected in linguistic processes. In what follows, the topics are taken directly from the psycholinguistic literature, and sentence processing mechanisms are discussed in the light of NCB computations, with emphasis on comprehension. Given that NCB language research is still in its early infancy (sections 1.3 and 2.3.1), the appropriate tasks allowing insights into linguistic emulation are yet to be integrated in NCB clinical and imaging studies. Thus, discussion will often be made on the basis of only indirect support from the findings available. Above all, a common thread running through these topics will be the striking similarities between the computations underlying the psycholinguistic mechanisms discussed and the NCB computational properties (section 2.2), as incorporated in the NCBKFLP model above (sections 2.3.3 and 2.3.4).

[24] See section 2.6.3.1 for further discussion in the contexts of the experimental hypotheses generated.

2.4.4.1. Pseudosyntactic Computations and Neocerebellar Output

The distinction between categorical and associative computations, mapped in the previous discussion to the properties of cortico-cortical and cortico-cerebellar processing, respectively (section 2.3.4.3), is often found in the sentence comprehension literature. According to the Late Assignment of Syntax Theory (LAST), for instance, sentence comprehension involves two different mechanisms that correspond to independent computations and representational levels: an ‘‘algorithmic’’ syntax, that employs ‘‘categorical’’ computations and reflects processing ‘‘rules’’, and a pre-emptive, ‘‘heuristic’’ ‘‘pseudosyntax’’, that employs ‘‘associative’’ computations and reflects processing ‘‘habits’’[25] (Townsend & Bever, 2001). Low-level cues, such as grammatical morphology, provide pseudosyntax the means to segment the sentential input into major constituents, assign probable thematic (θ)-roles (Chomsky, 1981b), and form an initial meaning-form hypothesis. The syntax subsequently uses this hypothesis to synthesize a detailed syntactic structure (Townsend & Bever, 2001).

The ‘‘rules/ habits’’ distinction invoked in such two-stage models meets the one often made in cognitive neuroscience between ‘memories and habits’. In the case of the ‘‘dual-systems model of retention’’, for instance, the memory system is instantiated in cortical and limbic loci, whereas the habit system is instantiated in basal ganglionic and CB ones (Mishkin & Petri, 1984; Mishkin *et al.*, 1984). The unitary, multimodal nature of CB computations promotes such a structure as a promising candidate for the storage and adaptation of syntactic and semantic processing habits. Moreover, this dynamic definition of the first stage of processing (‘‘habits’’) presupposes, apparently, a process of ‘habituation’, in other words, some adaptive mechanism whereby reoccurring processes in the second stage may be demoted to the first stage as a result of learning

[25] However, commitment to a two-stage model is far from necessary. See discussion in the end of this section.

(Argyropoulos, 2008b, June). The capacity of CB circuitry to adaptively copy the dynamics of repeated processes of other parts of the brain promotes, once again, the CB as a plausible neuroanatomical candidate for the instantiation of certain first-stage operations and such adaptive demotions to that stage. In other words, if a certain second-stage process for certain input reliably delivers a certain output, the NCB may acquire the dynamics of this process. Training signals would be derived from the discrepancies between NCB predictions for the output of that process (transmitted to the language-related cortical loci) and the actual output of that process (transmitted back to the NCB). The pre-emptive processing capacities of the CB (section 2.3.4) would then provide sentence comprehension with a number of pseudosyntactic mechanisms.

Interestingly, in some of these 'analysis-by-synthesis' models, production mechanisms are implicitly (Townsend & Bever, 2001) or explicitly assumed (e.g. Garrett, 2000) to participate in the first stage of sentence comprehension. In that stage, a sentence structure would be generated by the production system, and would shape the hypothesis about the sentence under comprehension. This type of process is implied in the ''scrambled-egg analogy'' below:

> […] the analysis-by-synthesis model starts with a particular hypothetical egg sequence, scrambles and cooks them in a virtual kitchen, and then compares the resulting virtual omelet with the actual input. When the virtual omelet matches the actual omelet, the input and cooking sequence producing the virtual omelet is confirmed as the correct analysis (Townsend & Bever, 2001, pp. 160-1).

This ''virtual kitchen'' here would arguably be accommodated by the CB, the brain's ''virtual reality machine'' (Miall, 2007). Moreover, the perception of actions produced by conspecifics has been argued to employ the perceiver's own CB internal

models (Blakemore & Decety, 2001),[26] which are used for the prediction of the consequences of the perceiver's own actions (e.g. Miall & Wolpert, 1996; Blakemore *et al.*, 2001; section 2.2.2 here). Again, then, this would provide further grounds for the involvement of CB computations in processes occurring at such a first stage in sentence comprehension.

Another aspect bringing the associative cortico-cerebellar linguistic computations proposed above closer to the associative pseudosyntactic ones here is the ''good-enough'' nature of the representations that the latter culminate in (Ferreira *et al.*, 2002; see below). In independent work in cognitive science, such 'good-enough' models of the perceived environment have been considered to confer significant advantages to their bearers:

> [n]aturally, the perceptual world is not totally out of synch with the 'real world'. The perceptual systems have evolved in order that organisms may *act reliably in the real world*. They are not concerned with a *'true model of the world' in the logical sense, but with a 'world model' good enough to support the planning of actions* that in the long run lead to better propagation of the genes**.** Like other products of evolution, the perceptual systems are full of 'cheap tricks', which is why we see virtual objects: these tricks work in the organism's normal environment. *It is only in the context of the laboratory that their artificiality is detected* (Jackendoff, 2002, p. 308; *italics* mine).

[26] This idea, developed in fact prior to Wilson and Knoblich (2005), could fruitfully provide grounds for arguing for a central CB role in emulative covert imitation. However, imitation processes are altogether beyond the scope of the investigation here. See section 2.5 for this point.

For the organisms to ‘‘act reliably in the real world’’ (*ibid*), then, the CB was seen above as fundamental to generating predictions for the sensory outcomes of actions, compensating for the conduction delays in updating the status of the peripheral motor system, enhancing perception overall on the basis of predictions on the perceived signal, or distinguishing reafference from exafference (sections 2.2.1 and 2.2.2). Furthermore, examples can be found in the literature on this ‘‘artificiality’’ (*ibid*) of CB-shaped percepts (section 2.4.4.6 below). Once again, then, the significance of CB computations in such action-oriented efficient perception, their domain-general nature, and the fact that this good-enough nature of perception processes is also reflected in those of sentence comprehension should all invite a closer look at the involvement of the NCB in constructing good-enough representations in sentence comprehension.

In language processing, such representations are often owed to the rapid decay of structure in working memory, and/or, crucially, to the susceptibility of the final analysis to influence from heuristic computations (Ferreira *et al.*, 2002). An important aspect of the latter is the idea that, in certain situations, the algorithmically-based analysis may not be obtained at all. This may be due to time pressure, or because the pseudoparse may have reached some ‘‘criterion level of confidence’’, terminating its operations (Ferreira, 2003, p. 171). Similar behavior is reproduced by the internal model in a Kalman filter. In sensory processing, for instance, Kalman filters are described to

> […] converge until they are weakly coupled to the target. At this point they tend to discount sense data more than they should because their computations are based on the assumption that the internal model is correct. Because the internal model can never be perfect, the estimator will converge to give suboptimal performance and may diverge from the target (Paulin, 1996, p. 530).

In sentence processing, this ‘weak coupling’ of the CB internal model with the copied target system can be seen as an aspect of such a criterion level of confidence being reached. Similarly, just as the internal model can never be perfect, so ‘‘[i]t is only the algorithmic parse that is guaranteed to yield a correct analysis’’ (Ferreira, 2003).

Interestingly, Deacon’s (1997) neurocognitive approach to the semiotic divide between ‘symbols’ and ‘indices’ reflects a similar perspective on the underlying foundation of this ‘syntactic/pseudosyntactic’ distinction: subcortically localized indices may help automate cortically distributed rules. Within a full fledged psycholinguistic model, such ‘‘indices’’ would correspond to ‘‘surface cues’’ triggering surface schemata, assigning θ-roles and separating major phrases (Townsend & Bever, 2001):

> […] grammatical cues, such as are embodied in small ‘function words’, may be the primary agents for initially tagging and distributing sentence ‘chunks’ to be separately processed. For this reason, it is precisely these features of language that need to be subject to minimal symbolic analysis. They serve a predominantly indexical function. And […] indices can be interpreted in isolation as automated, rote-learned skills. […] So, automatization of speech production and comprehension is accomplished by setting aside a small, closed set of symbols to be used as though they were indices. As a result of being stripped of semantic links, they can be learned by rote and implemented with minimal mnemonic search. And their representation within the brain can be highly local, even subcortical. The function of these modular operations is to implement grammatical rules, but the rules are implicitly symbolic and therefore distributed. These automated language functions are not grammar modules, but merely symbols of the grammar, which is itself probably highly distributed (Deacon, 1997, p. 299).

At this point, two clarifications are in order before continuing. Emulation in language processing is not proposed to exhaust pseudosyntax, but, rather, to only participate in such a stage. Correspondingly, from a neurological standpoint, it is not suggested that the NCB fully instantiates the pseudosyntactic component, but, rather, that it constitutes at least a fundamental site in what would conceivably be a circuit of interconnected brain loci (section 2.4.4.4).

Most importantly, however, it is necessary not to commit the above reflections exclusively to a constraint-based or a two-stage model of sentence processing. The latter may often receive attention in exemplifying CB contributions as opposed to cortical ones. However, this is done so as to illustrate the qualitatively different computations that are maximally involved in cases of minimum NCBKFLP gain. In other words, commitment to different computations here does not necessitate commitment to a particular fashion in which these should be temporally arranged (e.g. whether one computation operates in parallel with, or pre-empts the other).[27] On the one hand, for instance, the radically pre-emptive capacities of the CB (section 2.3.4) and of internal models per se in sensorimotor control (section 2.2.1) make the idea of a 'CB first stage' an intuitively satisfactory one. On the other hand, of course, the speed of CB processing would still not exhaust its role to a pre-emptive one: the Kalman filter is, in its own right, a stochastic state estimating formalism. This would mean that the NCBKFLP negotiates the biases of internal model outputs on the ultimate representations in language perception and comprehension. Multiple such biases of varying strengths would come to play, much as in constraint-based models. Thus, the extent to which constraint-based and two-stage models of sentence comprehension are compatible with the NCBKFLP is beyond the scope of the present thesis, and should be a topic for further investigation.

[27] The Pickering and Garrod (2007) processor, for instance, combines two different types of computations (perceptual/ comprehensional input analysis, and production-based emulation) without committing to a two-stage model.

2.4.4.2. Cerebellar Constrained Search

A major contribution of associative computations is that they minimize the search space in processing. Language comprehension-perception is often performed in time-critical and/or noisy conditions (e.g. Ferreira, 2003; Pickering & Garrod, 2007), and, for this reason, those constrained search mechanisms become of added significance. In some two-stage, analysis-by-synthesis models, for instance, the heuristics used in the first stage, such as semantic associations and the N-V-N strategy, limit the hypotheses ‘‘to sentences that have the indicated words with just those underlying structure roles that typically correspond to the conceptual θ-roles’’ (Townsend & Bever, 2001, p. 164). In other words,

> [t]he pseudo-parser is the component that uses heuristics […] to create a preliminary hypothesis. The true parser then uses that preliminary hypothesis *to constrain its search space*, so that when it calls up its syntactic procedures it need only consider a small subset of the formal possibilities (Ferreira, 2003, p. 170; *italics* mine).

While the specification of the time-course of such particular constrained search computations is beyond the scope of the investigation here, a NCB role in performing these computations can be defended on the following grounds. As discussed above (sections 2.2 and 2.3), CB emulation may function as a ‘‘look-up table’’ of past input/output instances of the emulated activity. For every relevant action, the system creates a new ‘‘entry’’ consisting of the initial state of the target system, the actions performed, and the final state. Upon receipt of the initial state specification and the efference copy of the command, the table searches for a match in its list of stored input-output pairs, and, upon retrieval, produces the ‘‘final state’’ specification on that entry as output. Such an implementation, of course, would constrain the generalization to new instances (Grush, 2004; Ramnani, 2006; also section 2.3.4.3 here). Purkinje cells in the

CB cortex, in view of the massive synaptic input they receive by granule cells, are often considered to play such role (Thach, 1996a). Similarly, in Doya's (1999) approach,

> […] a cerebellar mapping can work as a *short-cut circuit* or a *look-up table* for a mapping that was originally developed by the *time-consuming cortico-cortical processing*. Such a mechanism is especially useful when the *same mapping is used repeatedly* for *time critical tasks* […] (*ibid*, p. 970; *italics* mine).

Interestingly, imaging and clinical evidence has implicated NCB circuitry in constrained search in language processing, independently, though, from the computational literature on the CB. For instance, increases in activation in left-frontal cortex and NCB were found in an fMRI word-stem completion task (Desmond *et al.*, 1998). Stems with many (MANY condition; e.g. 'STA_') or few possible completions (FEW condition; e.g. 'PSA_') were presented, and subjects were asked to covertly complete each stem with a word. Prominent increases in activation in the MANY (relative to the FEW) condition were observed in the left middle and inferior frontal gyri (BA 9/10/46), left anterior cingulate (BA 24/32) and left caudate nucleus. In contrast, right NCB vermal and hemispheric lobules showed increases in the FEW (relative to the MANY) condition. Desmond *et al.* (1998) interpret this double dissociation on the basis of distinctive contributions of the two areas to cognitive performance: left-frontal and striatal activations reflect response selection, which increases in difficulty when there are many appropriate responses; NCB activation reflects search for responses, which increases in difficulty when even a single appropriate response is hard to retrieve. However, in the task involved, the FEW and MANY conditions were not matched for difficulty, as the FEW condition involved longer response latencies. Thus, for Desmond *et al.* (1998), NCB involvement in the FEW condition could conceivably reflect increased working memory demands associated with articulatory rehearsal.

The above findings have been replicated in a clinical study for a schizophrenic population (Marvel *et al.,* 2004), where NCB, especially NCB vermal pathology is predominantly involved (see section 5.4.7.2.1 here for references). The results demonstrated a selective deficit for schizophrenics in producing stem completions in the FEW condition. Contrary to the Desmond *et al.* (1998) study, the two conditions here were matched. In fact, latencies for control subjects were shorter in the FEW than in the MANY condition, suggesting that the FEW condition did not involve an increased demand for working memory. Moreover, working memory ability for the patients was not significantly correlated to their impaired performance in the FEW condition. For Marvel *et al.* (2004), these findings suggest that working memory does not fully explain the patients' deficits and that the NCB role in word search in schizophrenia may not be related specifically to working memory processes and articulatory rehearsal. This would add further weight to the interpretation here, namely, that the NCB is involved in constrained search in language processing.

A recent study on NCB patients has explicitly considered the size of the search space as a factor affecting patients' performance (Gebhart *et al.,* 2002). The range of appropriate responses in the antonym- (e.g. stimulus: 'KIND', response: 'rude') and verb-generation tasks reported (e.g. stimulus: 'CHEF', response: 'cook'), on which patients showed deficits, was much more constrained than that in the subordinate term-generation task (e.g. stimulus: 'MONEY', response: 'dollar'), where they showed spared performance. This pattern encouraged the interpretation that ''the element of constraint in semantic association responses may be a key factor in eliciting a right CB language deficit" (*ibid*, p. 332).

In the light of the NCBKFLP (section 2.3.4), stem completion in the FEW condition and word-generation tasks with equally few appropriate responses (e.g. verb generation) involve unification in a limited search space. The NCB would provide such a limited search space for re-occurring, routine operations, thereby attenuating the categorical search of the temporo-parietal memory component. There is, moreover, no

reason to limit NCB involvement to stems or simple lexical items. NCB constrained search would conceivably extend to higher-order constructional templates, ranging from lexical morphemes to particular constructions for unification. The homogeneity of CB architecture and its projections to the areas undertaking such unification operations adds weight to such a hypothesis. These NCB constrained-search linguistic processes are proposed in the next chapter to be of added significance in motivating grammaticalization phenomena (section 3.6.2). Sections 2.4.4.3 and 2.4.4.5 below discuss findings in the light of NCB constrained search in sentence processing.

2.4.4.3. Processing Passives without the Cerebellum

As mentioned above, research in NCB language functions is both in its early infancy and is heavily concerned with the frontal-like language symptoms that NCB damage may occasionally result in. The work in Pickett (1998), however, stands out as a characteristic case of CB deficits in sentence processing that cannot be explained on the basis of cortical-like patterns. Unsurprisingly, those findings have attracted very little attention,[28] and have not had any further impact on CB language pathology.

Pickett (1998) provides the first report in the literature of CB impairments in sentence comprehension (*ibid*, p. 103). In the two tasks reported, participants listened to the stimulus sentence (a description of a simple transitive event), and had to select the correct picture corresponding to the event described. The dependent measure was error rates. In the first task, where active and passive sentences were contrasted, normal control subjects made errors in 4% of the actives, and in 10% of the passives. On the contrary, CB patients made errors in 11% of both actives and passives, strikingly showing, unlike normal controls, no effect of Voice manipulation. Interestingly, the two patients with right NCB damage of the seven total patients performed better than normal control subjects.

[28] A 'Google Scholar' search (25 May 2010) shows but 5 citations of the work.

From a ‘frontal-like’ perspective, such lack of Voice effect was seen as ‘‘difficult to interpret’’ (Pickett, 1998, p. 91), ‘‘disappointing’’ (*ibid*, p. 90), and as providing ‘‘mixed results’’ (Justus, 2004, p. 1128). The pattern expected was that the more complex and difficult-to-process passive constructions would be the ones to show the impairment in the patient population. In such a framework, though, the lack of Voice effect would remain unexplained.

This pattern was difficult to attribute to a general cognitive impairment: no subject had below normal performance on the cognitive task (‘Digit Span Backwards’), the means and ranges on which were the most closely matched of the tasks for the two groups. This suggests, for Pickett (1998), that the two subject groups had roughly comparable cognitive-intellectual levels. Furthermore, in the second task, which compared simple sentences (‘The fireman kicks the chair’), sentences with sentence-final (‘The fireman kicks the chair that is soft’), and center-embedded relative clauses (‘The fireman that is fat kicks the chair’), the patients showed the expected effects of structural complexity, with higher error rates on more complex structures. Patients also showed the expected effects of semantic constraint like normal controls in both tasks: they made more mistakes in semantically unconstrained sentences (equivalent to Ferreira’s (2003) ‘‘symmetrical’’ ones; e.g. ‘The sister hugged the brother’), than in semantically constrained ones, where the two θ-roles could not be reversely assigned to the two nouns involved (like Ferreira’s (2003) ‘‘nonreversible’’ ones; e.g. ‘The child pulled the wagon’). [29] Patients were thus able to recruit semantic information in sentence comprehension. These findings suggested that the CB is not implicated in the processing of syntactic rules or lexical and sentential semantics.

[29] A closer look at the materials for both experiments reveals that all the stimuli manipulated belonged either to the equivalent of the ‘‘symmetrical’’ or of the ‘‘non-reversible’’ cases. In other words, unlike Ferreira’s (2003) stimulus sets, there were no ‘‘biased reversible’’ sentences (e.g. ‘The dog was bitten by the man’). This last type is the one where semantic associations heavily invite schematic interference in processing (section 2.4.4.5). Thus, NCB computations can be discussed here only in the light of the ‘N-V-N’ heuristic.

In view of such unexpected results, Pickett (1998) speculates that the voicing distinction may be associated with an automaticity distinction; active sentences, being more frequently employed than passives, would allow more automatic processing:

> […] the cerebellars do not find the passive sentences difficult, but […] they find even active sentences difficult. For example, it could be the case that active sentences normally are processed much more *automatically* than passive, while passive sentences may require *more attention* or perhaps more of a *word by word* rather than *phrasal parsing*. If the CB subjects are impaired, they may use the same more effortful strategy on all sentences and thus do not benefit from the relative simplicity of the active sentences […] So perhaps the linguistically impaired cerebellar subjects […] can perform the *full range* of linguistic processing, but their *mental shortcuts* have been disrupted, so that processing 'easy' sentences requires more conscious effort and is performed less automatically (*ibid*, pp. 91-103; *italics* mine).

Pickett's (1998) considerations above on the ''phrasal'', more ''automatic'' parsing of active sentences, as opposed to the more ''word by word'', ''attended'' parsing of passives, presuppose and implicitly promote a 'rules/habits' distinction. Such a divide is explicitly assumed in two-stage models of sentence comprehension discussed above (section 2.4.4.1).[30] Characteristically, one of the important tenets in LAST is that word sequences that conform to canonical sentence templates are easier to comprehend. For example, in sentence acceptability-judgment tasks (e.g. Forster & Olbrei, 1973), active sentences (e.g. 'the boy liked the girl') receive judgments at a fast rate, as opposed to simple passives (e.g. 'The girl was liked by the boy'). This is partly

[30] Once again, the 'rules-habits' distinction adopted here need not commit to a particular temporal arrangement of these two types of computations. Thus, in a constraint-based model, similar computations could be represented as competing biases of different strengths (see also section 2.4.4.1).

accounted for by the N(oun)-V(erb)-N(oun) heuristic that pseudosyntax applies (Townsend & Bever, 2001; Ferreira, 2003), according to which N-V-N sequences bias interpretation in favor of an 'agent-action-patient' output. For Pickett's (1998) CB patients, then, it seems that the ''cost for the passive structure overall'' (Ferreira, 2003, p. 177) becomes tantamount to that for the active structure. This would be due to the abolishment of a CB-based N-V-N heuristic, which otherwise guarantees fast and accurate processing of active voice sentences.

In particular, CB patients may show an 'advantageous inaccessibility' of the internal models in the NCBKFLP. In a healthy CB, the highly entrenched θ-role assignment of [Agent = N_1, Action = V, Theme = N_2] would be transmitted by the neodentate in response to efference copies of sentential inputs of the form of [N_1V N_2] sent from the cortical input analysis system to the NCB cortex via the cerebro-ponto-cerebellar pathway. Disruption of internal models copying θ-role assignment in these sentences would not result in impairments in processing their syntactic and semantic aspects, as it did not for Pickett's CB patients above. This would be because, in many tasks, CB contributions are of an optimizing yet unnecessary nature, and non-CB strategies might in fact be more efficient. Conceivably, ''[a] CB patient would be unable to use the CB-based strategy, but may be better at solving the problem'' (Paulin, 1997, p. 531). The non-CB strategy developed here in processing active structures may have made the two right NCB patients better than normal controls, and the rest of the (more cognitively impaired) patients indifferent to the active/passive distinction.

2.4.4.4. The Cerebellum, Broca's Area, and Pseudosyntax

As discussed above, the NCBKFLP is not committed to a two-stage or a constrained-based model of sentence comprehension (section 2.4.4.1). Instead, both types provide insights into the properties of NCB computations in language processing. In this section, agrammatic aphasic[31] sentence comprehension is discussed in the light of a spared pseudosyntactic component that may partly employ NCB circuitry.

The distinction between ''rules'' and ''habits'' (Townsend & Bever, 2001), or between ''algorithms'' and ''heuristics'' (Ferreira, 2003) has been predated (and inspired) by that introduced in aphasiology to account for the performance patterns of Broca's agrammatic aphasics (Caramazza & Zurif, 1976; Saffran *et al.*, 1998). The latter seem to operate in sentence comprehension based on default patterns, such as the N-V-N strategy, or semantic associations among lexical items in sentences (''schema-transmission''; Ferreira & Stacey, 2000). In fact, the performance patterns of Pickett's (1998) CB patients are the very opposite of those for both normal control subjects (Ferreira & Stacey, 2000; Ferreira, 2003), and Broca's agrammatic aphasics in comprehending passive sentences (e.g. Saffran *et al.*, 1998). In sentence comprehension tasks, both healthy normal controls, and, to a greater extent, agrammatic patients have shown high error rates on sentences with semantic constraints conflicting with the syntactically-based assignments (e.g. 'The hunter was shot by the deer'). The explanation proposed is that θ-transmission in passives produces a weaker (for normal controls, and impaired for agrammatic aphasics) binding between θ-role and syntactic positions. The resulting interpretation is thus vulnerable to schematic influence from semantic associations (e.g. the routine association of hunters with the act of shooting, and that of deer with the patient role; Saffran *et al.*, 1998; Ferreira & Stacey, 2000; Ferreira, 2003). The pattern of aphasics exhibiting the qualitatively same and

[31] Since no particular distinctions are discussed between 'agrammatic aphasia' and 'Broca's aphasia', the two terms are used here interchangeably.

quantitatively exaggerated performance patterns is suggestive of the fact that such semantic associations influencing θ-role assignment are preserved in Broca's agrammatic aphasia, and are thus not computed by Broca's area.

Assuming a different perspective, Townsend & Bever (2001) speculate that pseudosyntactic operations are instantiated in Broca's area, and dismiss the N-V-N pattern that Broca's aphasics reproduce as post-comprehensional strategic reasoning. However, there are a number of issues with this interpretation. If such pseudosyntactic component is (fully) instantiated in Broca's area, then it would be difficult to explain how a patient population other than Broca's aphasics exhibit selective deficits in such N-V-N heuristics, and not in syntactic and semantic aspects of sentence comprehension. This is precisely the case observed in Pickett's (1998) CB group above. This pattern, the very reverse of which is found in agrammatic aphasics (Saffran *et al.*, 1998) certainly poses a problem for the instantiation of pseudosyntax in Broca's area.

Townsend and Bever's (2001) proposal heavily relies on the findings of Linebarger *et al.* (1983) on the preservation of grammatical knowledge in Broca's aphasics: allegedly, these patients cannot understand sentences in θ-role- or plausibility-judgment tasks, yet can judge whether or not these sentences are grammatical. This suggests, for Townsend and Bever (2001), that Broca's aphasics have lost the ability to use pseudosyntax to develop an initial meaning, but have not lost the full grammar. This finding, however, is not commonly accepted. Other experiments have shown that agrammatic aphasics exhibit severe deficits in grammaticality judgment tasks involving passives.[32] This shows then that the agrammatic deficit depends on syntactic structure, and not on the linguistic task involved, e.g. grammaticality judgment vs θ-role/plausibility-judgment tasks (Grodzinsky & Finkel, 1998).

[32] In Grodzinsky and Finkel (1998), the deficit is observed in a larger set of structures, i.e. those with dependencies between traces and their antecedents. The discussion here need not commit to the particular interpretation provided in that study for the deficits observed.

Another counterargument could be made on the basis of lexical priming studies on Broca's aphasics. If the pseudosyntactic schemata (most fundamentally, semantic associations) are applied by Broca's area, then deficits should be found in associative priming in agrammatic aphasics. However, studies on 'associative' and 'semantic' priming in Broca's and Wernicke's aphasia show no impairments in automatically accessing lexical and/or semantic information (Hagoort, 1997). On the other hand, while no associative priming studies have been reported for CB populations, NCB patients have shown deficits in verb-generation, but spared performance in verb-selection tasks, along with evidence for spared semantic-categorical representations (see next section).[33]

Another issue would be the one-to-one identification of particular neuroanatomical loci with instantiations of psycholinguistic components. As already suggested, pseudosyntax need not employ computations of a single neuroanatomical locus (section 2.4.4.1). Rather, if Broca's area is to be viewed as a mechanism involved in pseudosyntactic processes, then it should do so by its participation in the cortico-cerebellar and cortico-striatal circuits that form the ''procedural'' component of language processing (Ullman, 2004).

In the context of two-stage models, then, the CB would be but one neuroanatomical structure supporting pseudosyntactic operations, such as schema-transmission and N-V-N heuristics. Some first clinical and imaging evidence on the CB would also suggest that this structure is not involved in processing grammar in a categorical, 'rules'-based fashion, thus further corroborating the cerebral-categorical/ CB-associative distinction made above (section 2.3.4.3) To begin with, findings from neuropsychological experiments employing artificial grammar learning paradigms suggest that Broca's area, and not the CB, is involved in the acquisition of grammatical structure. ''Artificial grammar learning'' (Reber, 1967) is a form of implicit rule-learning. After viewing a series of letter strings formed according to a finite state rule

[33] The third TMS study reported here (chapter 7) provides some first evidence for NCB involvement in schema-transmission.

system, subjects are capable of classifying new letter strings as to whether or not they are formed according to these grammatical rules, despite little conscious insight into the rule per se. A patient study employing artificial grammar learning has shown no implicit learning deficits in rule extraction for CB patients (Witt *et al.*, 2002). On the other hand, several studies, involving fMRI (e.g. Forkstam *et al.*, 2006; Petersson *et al.*, 2004, 2010), TMS (Udden *et al.*, 2008) and tDCS (de Vries *et al.*, 2010), or clinical data (Christiansen *et al.*, in press), implicate Broca's area in artificial grammar learning.

Furthermore, unlike left hemispheric perisylvian lesions, CB ones are not strongly related with acquired aphasias and dyslexias. This suggests that the CB is not integral to the access and representation of orthographic, phonological, semantic and syntactic information, but instead that it exerts a more indirect influence (Desmond & Fiez, 1998). Indeed, CB-induced aphasias are the exception rather than the rule in NCB pathology (Daum *et al.*, 2001), and often involve cortical deafferentization (section 1.3.3). And whenever they occur, CB language deficits are milder than classical aphasic syndromes, such as Broca's or transcortical motor aphasia, owing to the ''neurofunctionally redundant'' nature of representations stored there (Fabbro *et al.*, 2004). Pickett's (1998) findings above also suggest that syntactic and semantic representations are preserved in CB language pathology (section 2.4.4.3).

Suggestively, the CB patient of Fiez *et al.* (1992) failed to provide semantic-associatively appropriate verbs for given nouns in a verb-generation task (see next section for discussion); however, their responses were semantic-categorically related with the presented word stimuli. For instance, when asked to generate an appropriate verb for the noun 'pill', they responded with the adjective 'small', instead of the prominent for normal controls response 'take'; similarly, for 'dog', they responded with 'cat', instead of 'bark' (Fiez *et al.*, 1992, p. 161). This supports the idea that CB damage leaves semantic networks intact (Fiez & Raichle, 1997). In the same vein, the NCB patients in Gebhart *et al.* (2002) showed impairments in producing an appropriate verb in response to a noun (see section 2.4.4.5 below), but when asked to select the

appropriate verb among alternative ones for a particular noun, they showed spared performance. The patients were also able to produce appropriate subordinate term (e.g. 'apple')-responses to a superordinate term ('fruit') presented to them. These results suggested that lexical semantic representations were spared; they are, moreover, 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, for references).

Finally, the metasystemic nature of CB contributions to processes of other parts of the central nervous system (section 1.4) makes it ideal for the performance of such 'pseudosyntactic' operations: the latter function outwith the syntactic system, yet they are defined by what the latter performs routinely. To the best of my knowledge, no imaging or clinical study has ever sought an explicit dissociation between Broca's area and the CB with respect to such automatization processes. However, the decomposition of motor synergies, the deficits in conditioning, and the dual-tasking impairments accompanying CB pathology have been described as quite intrinsic to CB disorders, and have not been likened to any motor deficits involved in Broca's area pathology (see section 3.5.3). Similarly, CB lesions make performance in sensorimotor tasks only suboptimal. In movement control, CB lesions affect compound movements, but selectively spare those of single muscles/limbs, allowing patients to adopt a seriatim mode of execution for the more complex ones (sections 2.2.2 and 3.5.3). Even in the case of eye-blink conditioning, responses have been argued to remain relatively spared after CB lesions, but to be made largely suboptimal (Courchesne & Allen, 1997). This suggests that only a few tasks fundamentally rely on CB computations (Paulin, 1997; section 2.4.4.3 here).

The NCB may thus arguably instantiate a subset of associative computations observed in sentence comprehension. Further research in the field should attempt to establish the necessary dissociations among different structures and different operations in sentence processing.

2.4.4.5. Neocerebellar Schema-Transmission

Turning to semantic aspects of sentence comprehension, both the computational properties of the CB and some first suggestive neuropsychological evidence support the significance of the NCB in ‘‘schema-transmission’’ (Ferreira & Stacey, 2000). In two-stage models (section 2.4.4.1), schema-transmission represents a characteristic first-stage, heuristic process, involving the activation of event schemata upon encountering the relevant concepts (e.g. the association of a dog with the agent-role in a biting event; see Ferreira & Stacey, 2000; Ferreira, 2003; Townsend & Bever, 2001). Schema-transmission of course influences sentence processing across two-stage and constraint-based models (McRae *et al.*, 2005), and thus deserves closer attention.

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. While early work on schemata involved highly structured information stored in memory and accessed directly from it, more flexible mechanisms have recently been proposed: schemata are not the products of inferential manipulation of declaratively encoded representations, but are activated below the level of awareness and computed online in sentence comprehension (see McRae *et al.*, 2005 for discussion).[34]

The emulative properties of the CB are ideal to undertake the fast retrieval and online application of schemata. Despite the representation of events in long-term memory in temporo-parietal cortical areas, NCB ‘‘dynamics memory’’ (Ito, 1993a, b)

[34] The reader may notice that the notion of a ‘schema’ used in the sentence comprehension literature differs significantly from that used in the construction grammar framework (section 2.3.4.3). In the former, a ‘‘(constructional) schema’’ refers to abstract constructions, which are roughly equivalent to generativist rules (Diessel, 2004). In the latter, an ‘‘(event) schema’’ reflects real-world knowledge, and is close to the idea of a ‘script’ and action-based routines (e.g. Hurford, Flaherty, & Argyropoulos, 2007). This is an important clarification, as, in the former notion, a schema is discussed as a categorical, cortically-stored representation, whereas, in the latter, as involving CB expectancy-generation mechanisms. These quite confounding terminological contingencies are suggestive of the pervasive use of the term ‘‘schema’’ (Rumelhart, 1980) in cognitive science.

would guarantee their fast online application in sentence processing. For a sentence like ‘the dog was bitten by the man’, the NCBKFLP, in a low Kalman gain modality, may output an interpretation (dog′(x), man′(y), bite′(x, y)) based on the predicted, statistically prominent arrangement of the semantic representations of lexical items, whereby the dog occupies the agent position in a biting event. By doing so, the CB would compensate for the often noisy, ambiguous, or computationally demanding linguistic input by overriding and/or bypassing the slower cortical computations of the syntax/semantics interface – only the latter guarantee the correct, yet slower and infrequent interpretation (dog′(x), man′(y), bite′(y, x)).

Suggestively, priming phenomena observed in perceiving pairs of nouns (e.g. ‘dog’) and appropriate θ-related verbs (e.g. ‘biting’) have been accounted for in terms of ‘‘expectancy generation’’ (McRae *et al.*, 2005). As discussed above, the domain-general role of the CB in prediction and preparation for upcoming neural events (e.g. Courchesne & Allen, 1997) promotes it as an ideal candidate foundation for these mechanisms (section 2.2).

From a similar perspective, knowledge of ‘‘dynamic events’’ has been argued to impact θ-role assignment (Ferretti *et al.*, 2003). These events are construed as having a period that leads up to a change of state, a period in which this change occurs, and a period that follows. Entities associated with particular events become salient in particular time points of the event. For example, agents are typically associated with the initiating conditions and the ongoing event, as they tend to cause the event to occur (e.g. ‘The cop arrested…’). Patients are associated with the resultant states and the ongoing event, as they often undergo a change of state (e.g. ‘…arrested the crook’). As discussed above, CB feedforward control involves the acquisition of internal models for dynamic systems, i.e. systems the states of which change over time (section 2.2). If the perception of a particular entity is a reliable predictor of a change of state in an event (e.g. an agent), or vice versa (e.g. a patient), the low NCBKFLP gain conditions in processing such an event will invite maximal involvement of the NCB internal model. The latter

may override and/or bypass the perceptual-comprehensional input analysis system to provide a rapid arrangement of these entities and events into the scene in which they habitually occur (e.g. arrest´(x, y), x = cop´, y = crook´).

The strikingly identical nature of CB computations with those underlying the psycholinguistic mechanisms above makes it once again instructive to examine the neurolinguistic evidence available in the literature. Given that CB language research is still in its early infancy (sections 1.3.2, 2.3.1), it is hardly surprising that θ-role priming experiments have not been conducted on NCB patients. However, insights can be gained into the issue from the methodologically much easier to apply verb-generation tasks.

One of the first examples of CB activation during a clearly cognitive task came from a PET study reported by Petersen *et al.* (1989; also section 1.3.2 here), where subjects were asked to generate aloud semantically appropriate verbs in response to visually or aurally presented nouns (e.g. 'eat' in response to 'cake'). During the control scan, subjects were asked to read aloud or repeat auditorily presented nouns. Activation of an area within the right lateral CB was found when subjects generated verbs, but not when they read or repeated nouns. Thus CB activation was difficult to account for on a purely motor basis. Similar results have been found in a recent study, where activations of right posterolateral CB loci were shown to be rather selectively involved in a verb-generation task (Frings *et al.*, 2006; see section 4.3 for more discussion).

The first clinical case study to support a CB role in the selection of lexical concepts was that of Fiez *et al.* (1992). They examined an English-speaking patient with a large right CB infarct, who generated inappropriate responses in a number of word-generation tasks, involving verb generation. As discussed in section 2.4.4.4 above, their responses were semantic-associatively inappropriate, but remained semantic-categorically related (e.g. 'small', instead of 'take' or 'swallow', in response to 'pill'). This could not be attributed to overall cognitive impairment, since the patient's performance on tests of memory, intelligence, 'frontal function', and language skills was

excellent. Similar findings have been reported in subsequent studies (e.g. Fabbro *et al.*, 2000).

Stronger evidence is provided in Gebhart *et al.* (2002), where a striking dissociation is shown between spared performance of NCB patients in the verb selection task and pronounced difficulties in the verb-generation task. Patients performed poorly in generating appropriate verbs for a given noun, but were able to select the correct verb-response to a noun-stimulus from a list of alternative verb-responses. This suggests that the lexical semantic/syntactic representations of noun-verb associations were preserved in memory, however, the online ‘‘internal generation’’ of verb-responses to noun-stimuli was compromised (*ibid*). This dissociation between verb selection and verb generation could also be interpreted as a selective sparing of cerebral-stored lexical and syntactic information, along with a disruption in the CB-based online generation of habitual semantic associations. Furthermore, the dissociation between CB patients’ spared performance in subordinate term generation (stimulus: fruit; response: apple) and their impaired performance in verb generation suggests that ‘‘[t]he right posterolateral cerebellum may be more involved in associative semantics than in categorical semantics’’ (Gebhart *et al.*, 2002, p. 332). This would directly add weight to the model here, and to the distinction made between cortico-cortical, categorical processes, cortico-cerebellar, associative ones (section 2.3.4.3). Finally, the third TMS study reported here (chapter 7) provides some rather suggestive evidence in support of selective involvement of NCB circuitry in schema-transmission.

2.4.4.6. Neocerebellar Shallow Processing

Insights on the linguistic contributions of NCB computations can be gained by another central feature of sentence comprehension, namely, ‘‘shallow processing’’ (Sanford & Sturt, 2002) or ‘‘underspecification’’, whereby

> […] the comprehender rarely constructs a precise and complete semantic analysis because the information explicitly offered in the text is incomplete or because a precise complete interpretation would not be worth the processing effort (Sanford & Graesser, 2006, p. 100).

Cases of ‘‘pragmatic normalization’’ (Fillenbaum, 1974) and ‘‘verbal illusions’’ (Wason & Reich, 1979) provide some characteristic examples. The sentences in (1a) and (2a) are systematically misprocessed and assigned the plausible but not actual meanings of (1b) and (2b), respectively, instead of the consistent with respect to the syntax-semantic interface, yet absurd with respect to the semantics-pragmatics interface meanings along the lines of (1c) and (2c), respectively.

(1a) ‘Don’t print that or I won’t sue you’ (Fillenbaum, 1974)

(1b) If some item is printed, the result will be a lawsuit.

(1c) If some item is printed, the result will not be a lawsuit.

(2a) ‘No head injury is too trivial to be ignored’ (Wason & Reich, 1979)

(2b) Every head injury should be treated, no matter how trivial

(2c) Only trivial injuries should be treated**.**

These cases reflect the application of an interpretational routine triggered by the construction ‘don’t X or Y’ (1) and ‘no X is too Y to be Z-ed’ (2), respectively. They thus involve a top-down effect from pragmatics to semantics that overrides the computations of the syntax-semantics interface necessary for arriving at the correct interpretation. In processing these sentences, then,

> [i]t is likely […] that the plausible meaning may suppress the need for verification of the syntactic and semantic relation in the sentences by the subjects, beyond a certain point, simply because they cannot see the reason for such verification (Natsopoulos, 1985, p. 390).

For these illusions, the Pickering and Garrod (2007) production-based emulator has been considered to override the processes of the syntax-semantics interface in the input analysis system and provide an interpretation based on the habitual arrangement of the semantic components of these sentences. Verbal illusions thus reflect defining features of NCB feedforward control and state estimation. Reoccurring associations between inputs and outputs of certain processes (stored in the CB) may override and/or bypass these very (cerebral) processes, biasing or altogether tramping their 'actual' output, and inhibiting further analysis of their input (sections 2.2.2 and 2.3.4.3).

Further insight may be gained from another perspective, whereby verbal illusions are outputs of incorrectly constructed ''mental models''[35] (Garnham & Oakhill, 1996). As mentioned above, one of the most insightful considerations so far on CB internal model contributions to thought processing has been the idea that the NCB copies the dynamics of the construction of mental models. The prefrontal association cortex acts as a controller on the temporo-parietal cortical areas, i.e. the controlled object. By routinization of a particular thought operation, NCB internal models may copy the mental model and allow the controller to operate on the NCB rather than on the temporo-parietal cortex. In that way, thought may occur rapidly and well below the level of awareness (e.g. Ito, 2000, 2008; sections 2.3.1 and 2.3.4 here).

[35] The idea that people rely on mental models is traced back to Kenneth Craik's (1943) suggestion that the mind develops ''small-scale models'' of reality to anticipate events. They are constructed from perception, imagination, or the comprehension of discourse, with varying degrees of abstractness.

In the case of the verbal illusion here, the NCB internal models would receive a copy of the sequences of the representations of the individual lexical items that the cortical input analysis system is processing to construct the particular mental model as an output. In the processes of the cortical input analysis system, the prefrontal controller may unify the individual conceptual representations of the lexical items activated in temporo-parietal loci according to the specifications of the syntactic structure of the sentential input. On the contrary, the internal models will output, via the dentato-thalamo-cortical tract, the statistically reliable, habitual arrangement of those conceptual representations that is routinely generated by cortical loci (section 1.2.4). Low Kalman gain conditions are established here both due to the strength of the association between lexical semantics and habitual interpretation, and due to the heavy computational demands posed by the syntactic structure of such inputs. These conditions invite maximized NCB involvement, with neocerebello-cortical computations overriding and/or bypassing the slower cortico-cortical ones.

‘‘Semantic illusions’’ provide another characteristic phenomenon involving underspecification and shallow processing. A series of different factors affect detection rates and thus the strength of the illusion (see Sanford & Graesser, 2006, for a review). For instance, in the famous ‘‘Moses illusion’’ (Erickson & Mattson, 1981), the illusion effect has been shown to reach a maximum when the critical impostor word is out of linguistic focus. This is because detection rates increase when reformulating the original sentence in (3a) as a focused cleft construction and directly probing its interpretation, as in (3b) below (Bredart & Modolo, 1988).

(3a) How many animals of each kind did Moses put on the ark?

(3b) It was Moses who put two of each kind of animal on the ark. True or False?

(Bredart & Modolo, 1988)

Crucially, a significant factor for the strength of anomaly detection in semantic illusions is the ‘‘the extent to which the scenario predicts the [anomalous] item’’ (Barton & Sanford, 1993, p. 477), since ‘‘[…] detectability is a function of scenario-based expectation’’ (*ibid*, p. 483). For instance, in the ‘survivors anomaly’ below (4), detection rates are significantly lower when the term ‘survivors’ is used in a plane crash scenario (4a), but are higher when the term is used in the context of a bicycle crash scenario (4b).

(4a) When an airplane crashes, where should the survivors be buried?

(4b) When a bicycle accident occurs, where should the survivors be buried?

(Barton & Sanford, 1993)

In (4a), low Kalman gain conditions are established by the strength of the associations between lexical semantics and routine interpretations. By maximized involvement of cerebro-cerebellar computations in determining the final output of the interpretational process, extensive cortico-cortical analysis of the properties of the linguistic input is inhibited. Just as low Kalman gain conditions in perception leave room for Kalman filters to over-discard sensory data on the basis of internal model predictions (section 2.4.4.1) so would the NCBKFLP here inhibit complete semantic analysis. Furthermore, as in the case of schema-transmission above (section 2.4.4.5), these ‘‘scenario-based expectations’’ here reflect CB computations of prediction and preparation, operating well beyond fully specified inputs and outputs:

> [c]omplete knowledge of upcoming events is not necessary to trigger specific preparatory actions. What is important is the probabilistic relevance of a particular sequence or stimulus as a predictor of an upcoming event (Courchesne & Allen, 1997, p. 3).

Suggestively, CB pathology has been associated with some striking violations of default economy principles in cognitive performance. For instance, when asked to write a sentence, a CB patient wrote eight lines; when asked to draw a clock with numbers on it, they drew two intricate diagrams, one of a wrist watch and one of a wooden tower clock (figure 2.6). This ‘‘cognitive overshoot’’ provides a characteristic case of the cognitive-affective abnormalities following CB pathology that fall under the term ‘‘dysmetria of thought’’ (Schmahmann, 1998). In general, these involve disruptions in the universal function of the CB ‘‘as an oscillation dampener, maintaining function steadily around a homeostatic baseline’’ (*ibid*, p. 367). On the basis of the above, then, pragmatic normalization would represent an instance of NCB oscillation dampening in sentence comprehension.

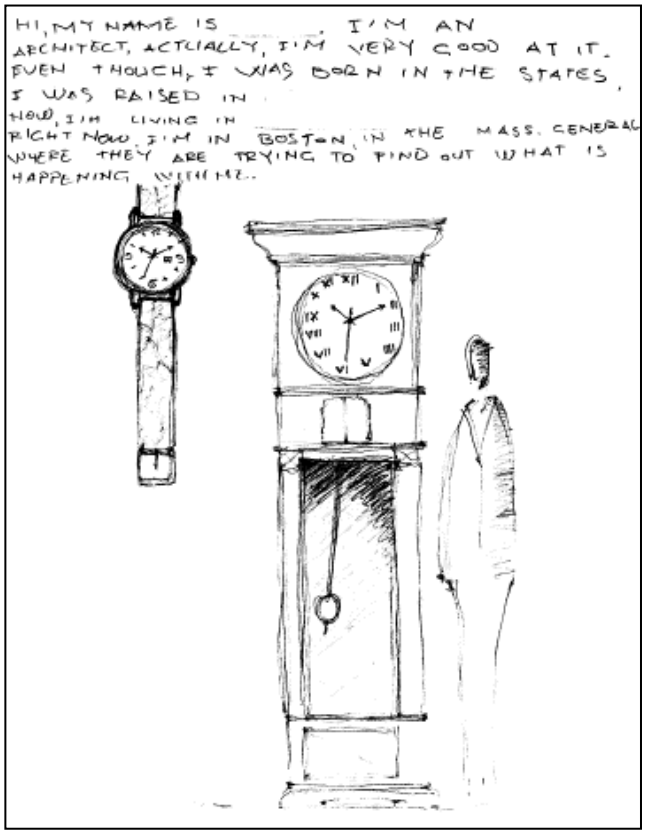


Figure 2.6: Responses of a CB patient to requests to write a sentence and draw a clock with numbers on it. Figure taken from Schmahmann (1998, p. 366) with permission © 1998 Elsevier.

On a more speculative basis, Deacon (1997) has suggested that the CB may undertake a ‘‘rapid’’ but ‘‘shallow’’ search for word associations, the results of which are subsequently selected by the prefrontal cortex. The CB may thus

> […] be very important for the generation of word associations at a rate sufficient for speech. At the rate words are presented in speech, the speaker or listener must be able *rapidly to generate associated words* and avoid letting earlier associations interfere. The cognitive search process must be *as rapid but as shallow as possible*. Any slight tendency to perseverate would entirely derail the process […] The cerebellum may provide *an independent generator of novel but predictable shifts of associations* from one context to another, while prefrontal cortex is providing a selective inhibition of all but the one sample that fits the new criteria (*ibid*, p. 277; *italics* mine).

This CB ‘‘shallow cognitive search’’ then may be at least partly accountable for underspecified semantic analyses. For instance, in the survivors anomaly above (4a), the CB-generated ‘‘predictable shifts of associations’’ from ‘airplane crash’ to ‘survivors’ would provide biasing factors for the illusion to occur.

Perhaps the strongest piece of suggestive evidence comes from independently motivated studies on the role of the CB in prediction-based illusions. The CB is involved, for instance, in the ‘‘size-weight illusion’’, whereby the smaller of two objects of equal weight is perceived as heavier (Hubbard & Ramachandran, 2004). This has been explained as a mismatch between predicted and actual sensorimotor feedback. Resistance to such illusions has been shown in both CB patients and schizophrenic populations. In Hubbard and Ramachandran (2004), normal control subjects and patients were initially tested for their accuracy in discriminating different weights. They were

presented with a pair of cans differing in weight by 50 grams and were asked to state which of the two was heavier. All subjects showed accurate weight discrimination in control conditions. The magnitude of the illusion was assessed by asking subjects to determine among the number of small cans of different weights which one matched the apparent weight of a large, heavy one. Control subjects showed clear effects of the illusion, matching the large can with one weighing substantially less. The illusion was not due to an inability to discriminate the weight of the cans, since the magnitude of the illusion was far greater than the minimum difference that could be discriminated. On the other hand, five of their seven CB patients showed a reduction of the illusion, despite intact weight discrimination. Similarly, the reduced size-weight illusion found for schizophrenic subjects in a more recent study has been discussed in the context of disrupted forward internal models involving the CB (Williams *et al.*, 2010).

To a certain extent, then, phenomena like verbal illusions can be fruitfully understood as the linguistic equivalents of this ‘‘predictive sensorimotor illusion’’ (Hubbard & Ramachandran, 2004, p. 407) here. The cytoarchitectural homogeneity of the CB, its reciprocal connectivity with language-related cortical loci, along with the evidence for resistance to such illusions because of CB disruptions, and the fact that the same basic predictive computations are employed in sentence comprehension, all add weight to the idea that NCB circuitry may be fundamentally involved in the sentence comprehension illusions above.

2.5. The Neocerebellum, Mirror Neurons, Emulation and Imitation

As already suggested, the accommodation of feedforward control and state estimation in cognitive science and psycholinguistics has progressed outwith the context of CB neuroscience (section 2.3.1). Instead, the mirror neuron system, and, for language processing, its compartments in BA 44/45, have received emphasis to a smaller (Wilson & Knoblich, 2005) or larger (e.g. Iacoboni, 2005; Hurley, 2008; Glenberg, 2009) extent as the candidate instantiation of such computations,[36] neglecting CB circuitry as their fundamental site (see discussion in Miall, 2003). The discussion turns briefly now to the relation of the NCBKFLP with imitation mechanisms and mirror neuron functions in language perception-comprehension.

To begin with, the idea of premotor mirror neurons instantiating feedforward control has recently been met with skepticism. For instance, it has been argued that the impetus for this idea is partly motivated by the emphasis placed on the perception of grasping actions. Grasping involves a multi-state operation, the production of which per se involves feedforward control. However, observing simple activities-states, e.g. holding an object, which do not necessitate these computations, also involves premotor mirror neuron activation. This then makes prediction mechanisms an unlikely function for such a locus (Rizzolatti, 2009).

Above all, however, the hypothesis here does not need to refute the role of mirror neurons or of any other structure in covert imitation in language perception-comprehension. This is because the desideratum of this thesis has been to investigate how CB feedforward control and state estimation may be manifested and explored in language processing, whether within or outwith the context of covert imitative production mechanisms. This is why the NCBKFLP is not proposed to directly

[36] The Pickering and Garrod (2007) model remains compatible with, but uncommitted to mirror neuron involvement (Pickering & Garrod, 2009b).

instantiate the mechanisms of Pickering & Garrod (2007) or of any other related cognitive model, but only to gain the necessary psycholinguistic insight from them. As a result, the question of production mechanisms and the extent to which they are involved in emulative language processing is almost entirely beyond the scope of this investigation.

Rather, if any insights may be afforded at all into the neural grounding of production-based emulative language perception-comprehension, the claim would be a rather conservative one; namely, that only its emulative component has NCB circuitry as the strongest neural candidate. The idea then that the emulator is provided by the production system at the psycholinguistic level (Pickering & Garrod, 2007) should perhaps not exhaust its neuroanatomical foundations to language production-specific loci. Despite the adaptive coupling of emulation and covert imitation in cognitive repertoires, the underlying neuroanatomical foundations of the two processes might in fact be distinct. Conceivably, in the imitative perception of conspecifics, language production-related cortical loci may employ the CB as the par excellence, modality-independent locus of emulation (section 2.2.2), by their reciprocal connectivity (section 1.3.2). In Argyropoulos (2009), for instance, the NCB and frontal cortical mirror neurons were explicitly proposed to provide the grounds for Pickering and Garrod's (2007) processor in toto. A cooperation of imitation and emulation mechanisms could be thus undertaken by the frontal cortical-NCB circuits respectively.

Indeed, the current literature provides some first suggestive evidence for the involvement of the CB in emulative imitation: While the properties of the connectivity of BA 44/45 mirror neurons with the right posterolateral CB have not yet been researched, CB afferents via the dentato-thalamic tract to the F5, i.e. Broca's homologue in monkeys (Petrides & Pandya, 1994) have been held to have input-output relations with the mirror neurons in that area, representing the ''desired state'' in tool-use learning (Imamizu *et al.*, 2003). The connectivity between mirror neurons in Broca's area and the

right posterolateral CB would thus guarantee the imitation-emulation cooperation proposed above.

Certain issues in dyslexia also lie at the intersection of the motor theory of speech perception and CB functions. In the former, the core premise is that speech perception involves access to the speech motor system, and that the efficient perception of the invariant events of the speech stream in the variable acoustic signal may employ connections between production and perception (Liberman *et al.*, 1967; Galantucci *et al.*, 2006). In this context, the phonological deficit in dyslexia has been argued to result from an articulatory problem, promoting the CB as a ‘‘neural instantiation of the motor theory of speech perception’’ (Ivry & Justus, 2001, p. 513). Future research could thus fruitfully combine insights from CB deficits in dyslexia (Nicolson *et al.*, 2001; section 3.3.5 here) with those provided by the motor theory of speech perception to investigate CB involvement in the production-perception links.

Some imaging findings also implicate CB circuitry in imitative processes. In a functional imaging study, subjects were scanned while they watched actions for later imitation. Activations were found in the CB, as well as in the same parietal and premotor regions that are involved in producing actions. These activations indicated that, during observation of action, the neural networks subserving motor representations are already tuned for imitation (Decety *et al.*, 1997). This would be expected if, during imitative action perception, the brain recruited the CB-stored associations of the perceiver’s own intentions with the same contexts and consequences for those actions (see Blakemore & Decety, 2001, for discussion).

In this respect, then, covert imitative language production in perception-comprehension may be fully compatible with the more general treatment of NCB emulation here. Future research would be required to dissociate the contributions of the language production-related cortical loci from those of the modality-independent NCB feedforward control and state estimation. The thesis now turns to a set of experimental predictions directly generated by the considerations in the sections above.

2.6. Experimental Predictions

The phenomena grounding the psycholinguistic mechanisms of interest here (section 2.4) may offer the testing grounds for the neurolinguistic hypothesis presented. The predictions below constitute an agenda much larger than the purposes of the thesis, and are thus bound to be part of future research (section 8.2).

At an elementary level, all predictions below refer to selective lesions on one of two basic CB sites: the deep CB nuclei or the CB cortex.[37] With a certain degree of abstraction, the dynamic trade-off in the strengths of the outputs of these two areas (section 1.2.3.1) would indicate that a lesion on one would make the output of the other prevail. Conceivably, a lesion of the deep CB nuclei should abolish transmission of the linguistic prediction to the cortical input analysis system. This would result in a 'bottom-up' processing of the stimuli, heavily based on the sensory and categorical properties of the input. Overall, this would be captured with lower amenability to prediction-based illusions for CB patients, as already observed in the sensory domain (section 2.4.4.6). On the contrary, a selective lesion in the CB cortex would result in insufficient inhibition of the deep CB nuclei (which are constantly excited by mossy fiber and climbing fiber collaterals; section 1.2.3) and would thus involve a suboptimal over-reliance of processing on predictive signals via the dentato-thalamic tract to the language-related cortical loci; this could translate into increased amenability to such illusions.

However, currently available evidence does not suffice to strongly commit to associations of particular lesions with a particular direction of an effect, as above. For instance, it remains unclear whether state estimates are expressed either in the CB cortex or in the nuclei or in both (see Miall & King, 2007, for discussion). Thus, the primary commitment of the predictions below pertains to the qualitatively different performance that CB patients are expected to show from healthy normal controls.

[37] Unfortunately, CB pathology rarely exhibits such selective lesions. See section 4.3 for discussion.

2.6.1. Abnormal[38] Sentence Comprehension

The cases of sentence comprehension discussed above (section 2.4.4) can all be studied in experimental settings, with abnormal performance resulting from a compromised (silenced or disinhibited) predictive, top-down processing modality.

2.6.1.1. Garden-Path Amenability

Given that ‘‘eager’’, ‘‘predictive’’ top-down parsing supports strong incrementality (see discussion in Sturt & Lombardo, 2005), and that strong incrementality underlies garden-path effects (e.g. Frazier & Rayner, 1982), the impairment of neural components of a top-down processing modality will arguably have an impact on the amenability to these effects. As discussed above, a low NCBKFLP gain modality would invite maximized NCB involvement: in such case, syntactic predictions and schematic knowledge may override a syntax-based θ-role assignment (sections 2.3 and 2.4). A lesioned neodentate and a subsequent disruption of a top-down processing modality would translate here into stronger resistance to garden-path effects in processing sentences like (5a-b) below. On the contrary, selective CB cortical lesions would induce disinhibited neodentate signalling to cerebral output loci, inviting an overflow of predictive signals, and thus allowing higher amenability to the incorrect parse. [39]

[38] The term ‘abnormal’ is preferred, because differences in the localization of the CB deficit may involve effects in opposite directions with respect to performance, either in the form of ‘enhancements’, or of ‘impairments’ as compared with normal control subjects.

[39] From another perspective, the ‘‘Intonation Unit storage hypothesis’’ predicts that constructions that are ‘‘precompiled’’ (not analytically computed) are the Grammatical Units that normally occur in a single Intonation Unit. The hypothesis further predicts that incorrect parses in garden-path sentences are made on the basis of such precompiled grammatical-intonation units (Croft, 1995). The significance of the NCB in processing such units was discussed above (section 2.3.4.3), and is also treated from another perspective in the next chapter (section 3.6.3.1.2).

(5a) The patient persuaded the doctor that he was having trouble with to leave

* [NP V NP that S´]
VS:
[NP V NP that RC VP]

(5b) I told the boy the dog bit Sue would help him.

* [NP V NP S´]
VS:
[NP V NP RC S´]

(Croft, 1995)

2.6.1.2. Misinterpretation of 'Non-Canonical' Sentences

Earlier in this chapter, suggestive evidence was discussed for NCB-induced impairments in schema-transmission and/or syntactic pseudoparses (sections 2.4.4.1, 2.4.4.3, and 2.4.4.5). Sentences often misinterpreted on the basis of such mechanisms by normal control subjects as in (6) below would receive a qualitatively different treatment by CB patients. A disrupted dentato-thalamic tract would leave the system to operate on the basis of the categorical properties of the sentence input, without the influence of such schematic information or pseudoparses. On the contrary, a selectively disrupted CB cortex would fail to inhibit the predictive output of the constantly excited neodentate, thus resulting in a superfluous employment of such pseudoparses and schematic knowledge, in the expense of the cortically-based analysis.

(6) 'The dog was bitten by the man' (Ferreira & Stacey, 2000; Ferreira, 2003).

2.6.1.3. Semantic Illusion Amenability

In the same vein, selective neodentate pathology would disrupt the ‘‘scenario-based predictions’’ that normally affect the depth at which sentences like (7) below are processed (section 2.4.4.6). In that way, such CB patients would show higher success rates than normal controls in detecting the anomaly. On the contrary, selective disruption of the CB cortex would maximize amenability to semantic illusions, because of the disinhibited transmission of such expectations by the neodentate.

(7) ‘When an airplane crashes, where should the survivors be buried?’

(Barton & Sanford, 1993)

2.6.1.4. Verbal Illusion amenability

Neodentate disruption is also expected to minimize reliance on the interpretational routines involved in verbal illusions (see section 2.4.4.6 here). The interpretational burden would instead be allocated to the slower yet more precise bottom-up processor of the syntax-semantics interface. This would translate into significantly higher success rates in paraphrasing sentences like (8). This would not result from any increased processing capacities per se for CB patients,[40] but would instead be owed to increased reliance on bottom-up processing, in the face of smaller interference of any interpretational routines. Selective pathology of the CB cortex, on the other hand, would induce maximized reliance on interpretational routines, and would thus be associated with lower accuracy rates than those of normal control subjects.

(8) ‘No head injury is too trivial to be ignored’ (Wason & Reich, 1979)

[40] In fact, the global cognitive impairments in CB patients might make such stimuli far from ideal for a sentence comprehension task (see section 4.3 for discussion).

2.6.2. Abnormal Lexical Associative Priming

As discussed in the beginning of this chapter, the CB is involved the prediction and preparation of the nervous system for a neural event B based on its reliable temporal contiguity with the detected event A (Courchesne & Allen, 1997; see section 2.1 here). In language processing, one of the most elementary phenomena in which such processes manifest themselves was discussed to be that of 'lexical associative priming': a prime word A facilitates recognition of target word B, given the frequent co-occurrence of their phonological and/or semantic representations in discourse (section 2.3.4.3).

2.6.2.1. Formal-Associative Priming

In the case of formal associates, as in the lexical pairs in (9) below, CB cortical disruptions would result in insufficient inhibition to the deep CB nuclear predictions. In such case, increased associative priming sizes would be a conceivable pattern in lexical decision task settings. In fact, this was the finding in the first TMS study reported here (chapter 5). On the other hand, disruptions in the dentato-thalamic tract would offer no room for prediction to influence processing at the cortical level. In such case, recognition of the temporal contiguity of the two items would rely on slower cortical processes. This would be reflected by smaller associative priming sizes.

(9) gift-HORSE, kidney-BEAN, livery-STABLE (e.g. Moss *et al*., 1995)

2.6.2.2. Semantic-Associative Priming

The same mechanism could be tested for semantic associations. Online application of schematic knowledge would involve NCB computations. Thus, encountering particular components (e.g. θ-role fillers) pertaining to ‘‘dynamic events’’ (e.g. referents of verbs) should trigger CB expectancy generation for this event to be processed (section 2.4.4.5). This is the case in the θ-related noun-verb pairs in (10) below. In a priming paradigm, a CB patient group with selective dentate pathology would show disrupted access to expectancy-generation mechanisms, demonstrating little facilitation in processing such pairs, and thus small priming sizes. On the contrary, selective CB cortical disruption would invite disinhibited associative predictions by the neodentate. Conceivably, the latter case could manifest itself in larger priming sizes for that particular type of priming. This pattern was observed in the third TMS study reported here (chapter 7).

(10) assassin-KILLING, merchant-SELLING, chef-COOKING (e.g. McRae *et al*., 2005)

2.6.3. Abnormal Phonological Predictions

2.6.3.1. Phoneme Restoration

Phoneme restoration reflects a phenomenon whereby the phonemes that correspond to the missing or noisy segments of perceived speech are restored by the hearers on the basis of expectations from the surrounding acoustic context. In other words, ‘‘[…] the stronger the contextual constraints, the greater the expectations should be, and the less bottom-up confirmation that should be needed’’ (Samuel, 1981, p. 476). In (11) below, for example, the speech sound removed was the first ‘s’ in the word ‘legislatures’, along with portions of the adjacent phonemes, and was replaced by a

cough. The sentential predictability of the word, as well as the intra-lexical, syllabic predictability, compromised the perception of the cough as a sound other than an [s].

(11) 'The state governors met with their respective legiSlatures convening in the capital city'

(Warren, 1970)

Phoneme restoration has already been discussed as reflecting emulative computations in speech perception (Pickering & Garrod, 2007). Evidence for the involvement of CB internal models in auditory enhancement in noisy contexts was presented in section 2.4.3 above. Conceivably, then, deep CB nuclear lesions would result in poor predictions on the upcoming speech segments. The percept would thus be analyzed solely on the basis of its sensory properties. In phoneme discrimination settings (Samuel, 1981), this would be reflected in elevated accuracy in detecting excised and replaced phonemes. In contrast, normal control subjects would perceptually enhance the noisy signal and thus restore it. Top-down factors, such as the number of syllables, the position of the selected phone, or the frequency of the word where such noise is introduced (*ibid*) would not affect the performance of those patients. On the contrary, selective CB cortical pathology would leave the NCBKFLP operating in a low Kalman gain modality. For such patients, the actual percept would be sub-optimally tramped by the disinhibited interference of sensory expectations, thus making the discrimination of a phoneme from the noise much more difficult.

2.6.3.2. Speech Reduction

As discussed in the beginning of the chapter, state estimation extends well beyond perception: for example, in contexts of low Kalman gain, it ensures a relatively feedback-free, rapid and accurate execution of limb movements in noisy and/or time-critical conditions (section 2.2.1.2). In the same vein, in phonetic-phonological processing, CB predictions are involved in speech production to the same extent as

speech perception. This is reflected in phenomena of articulatory reduction in cases where the planned items are predictable in their sentential contexts.[41] In other words,

> […] both the acoustic realization and auditory perception of a given word in a meaningful sentence may be a function of the speaker's and listener's knowledge of the semantic and grammatical information contained in the entire sentence **(**Lieberman, 1963, p. 172).

Independently of the accompanying dysarthrias (sections 1.3.1, 1.3.3), CB lesions might show selective impairments in speech reduction phenomena in a number of ways. Selective CB cortical lesions would result in disinhibited predictions in planning the articulation of a predictable lexical item as in (12a), as opposed to that of the same item in non-predictive contexts (12b). This may result in significantly more reduced articulatory gestures for the particular item in comparison with a normal control group in a sentence like (11a). On the contrary, silenced internal model predictions due to deep CB nuclear impairments would involve no facilitatory preactivation of the representation of the particular item in planning the corresponding speech segment. For such a patient group, no differences should be expected in the articulation of the item between sentence (12a) and (12b).

(12a) A stitch in time saves nine. (Lieberman, 1963)

(12b) The number that you will hear is nine.

[41] See also section 3.6.3.1 for a discussion of phonetic attrition in grammaticalization changes as involving CB-based articulatory reduction.

2.7. Conclusion

The lack of interdisciplinary discourse between research on CB feedforward control/state estimation and psycholinguistics has left the former in a stagnant status with respect to the involvement of these computations in language processing. Their recent introduction in independent work in psycholinguistics offers a unique opportunity to identify mechanisms the dynamics of which can be copied by NCB circuitry. In the context of the model proposed above, the first steps were taken to bring the two disciplines in fruitful discourse. In particular, neocerebello-cortical processes were argued to provide associative computations that may bias, override and/or bypass cortico-cortical categorical ones. A number of psycholinguistic mechanisms were shown to rely on computations the instantiation of which fundamentally involves CB circuitry. Clinical and imaging findings were also discussed in the light of this synthesis. The first experimental predictions were finally formulated, thus making the hypothesis a directly falsifiable one, and opening a broad research agenda for future work.

Chapter 3

The Neocerebellar Foundations of Grammaticalization

> ‘‘[O]ur ultimate goal is […] to uncover the actual mechanisms of change that operate in everyday language use that eventually give rise to grammatical categories’’ (Bybee *et al.*, 1994, p. 24).

3.1. Introduction

There are no ‘synchronic’ or ‘diachronic phenomena’; rather, there are different perspectives one may assume on the time-scale of different phenomena (Lehmann, 2004). In the same vein, the fact that grammaticalization[42] is predominantly studied from a diachronic perspective does not deprive the process from its synchronic aspect. In Argyropoulos (2008a), it was proposed that the neocerebellum (NCB) and the basal ganglia provide the neural grounding for the domain-general automatization phenomena that underlie the ontology of grammaticalization when examined from such a ‘synchronic’ perspective (see section 3.5.2 below, for references). It was further argued that current research on those structures allows the first steps towards constructing a neurolinguistics of grammaticalization. This is the line that I will expand on in this chapter. What I will argue in particular is that grammaticalization changes, at the intra-generational level of routinized adult dialogical interaction, show a gradual involvement of cerebellar (CB)-based associative computations, overriding and/or bypassing cortical, categorical ones in online processing. The Neocerebellar Kalman Filter Linguistic

[42]The term also appears in other forms, e.g. ‘grammaticization’, ‘grammaticisation’, ‘grammaticalisation’, or ‘grammaticalization’. The term used here is ‘grammaticalization’. See Hopper & Traugott (1993, pp. xv-xvi), and Lehmann (2002, pp. 8-10) for discussion.

Processor (NCBKFLP) is proposed to control the changes in the trade-off between those two computational types (section 2.3.4.3). After a brief description of the ontology of grammaticalization operations (section 3.2) and the underlying cognitive mechanisms (sections 3.3-3.5.2), I go on to show the fundamental importance of CB circuitry for such mechanisms (section 3.5.3), and discuss how the NCBKFLP could ground the particular neurolinguistic processes of grammaticalization (section 3.6). I finally discuss in brief the relation of other brain structures with this historical process, namely, that of the basal ganglia (section 3.7) and Broca's area (section 3.8).

3.2. Defining Grammaticalization

‘‘Grammaticalization’’, a term introduced by Meillet (1912), traditionally refers to ‘‘the increase of the range of a morpheme advancing from a lexical to a grammatical or from a less grammatical to a more grammatical status’’ (Kurylowicz, 1965, p. 52). Such definitions, nearly half a century later, remain at the core of standard grammaticalization theory: ‘‘[g]rammaticalization concerns the evolution from lexical to grammatical forms and from grammatical to even more grammatical forms’’ (Heine & Kuteva, 2002, p. 377).

3.2.1. Constructional Contexts

However, with the advent of construction grammar as a new promising paradigm for linguistic description, ‘‘[t]he constructional tail has come to wag the syntactic dog: everything from words to the most general syntactic and semantic rules can be represented as constructions’’ (Croft, 2001, p. 17; also section 2.3.4 here). Correspondingly, in historical linguistics, grammaticalization is considered to be more accurately defined as the process by which a lexical item becomes a grammatical morpheme only within the morphosyntactic context of a particular construction (e.g. Bybee, 2002a; Heine, 2003; Himmelmann, 2004, p. 31). In other words, it is constructions that become grammaticalized, and not individual lexical items. The grammaticalization of the latter is, in fact, ‘‘a by-product of the grammaticalization of a construction’’, and grammaticalization may occur without a particular item mediating between the constituents of a grammaticalizing construction (Lehmann, 2002, p. 7). This can be seen in the comparison between (1a) and (1b) below. The subjunctive in (1a) expresses the speaker’s distance from the content of his utterance, whereas in (1b) it is governed and triggered by the verb, and is thus more grammaticalized.

(1a) Erwin entschuldigte sich; er habe mich nicht absichtlich getroffen.
'Irvin apologized, he didn't hit me on purpose.'

(1b) Erwin sagte, er habe mich nicht absichtlich getroffen.
'Irvin said he didn't hit me on purpose.' (Lehmann, 2002, p. 7)

With the above in mind, it is instructive, at this point, to present a basic ontology of grammaticalization. The presentation is far from exhaustive of the range of phenomena involved, and is rather meant as a short list of selective highlights for the purposes of the investigation here.[43] Furthermore, the phenomena below cannot be identified completely beyond theoretical commitments, especially in the context of grammaticalization studies and the much heated debates often involved – in different frameworks, certain changes receive greater emphasis than others (e.g. sections 3.3.3 and 3.4.3).

3.2.2. Basic Ontology

Before the ultimate development of a new construction, grammaticalization changes operate on **particular instances of more schematic constructions**, i.e. constructions filled with particular lexical items. For instance, while several movement verbs instantiate the English constructional schema in (2), the only construction that has grammaticalized into a future marker is (2a), undergoing phonological, morphosyntactic, semantic, and pragmatic changes. Such changes isolate the particular construction from the generic one and its categorical coordinates in (2b) and (2c).

[43] For a thorough, parametric analysis of grammaticalization changes, see Lehmann (1995[1982])).

(2) [[movement verb + Progressive] + purpose clause (to + infinitive)]

(2a) I am going to see the king

(2b) I am travelling to see the king

(2c) I am riding to see the king (Bybee, 2003)

These multi-level changes that make the particular construction deviate from its corresponding schema include, among others, the following:

The grammatical items in grammaticalizing constructions are often derived from **metaphorically extended** lexical source concepts. Body parts, for example, on account of their relative location, are used as structural templates to express deictic location, as in (3) below. Temporal concepts often originate from prepositions expressing spatial relations, as in the transition from (4a) to (4b).

(3) dori- tutuo (Mursi)
house-mouth:of (Turton & Bender, 1976, p. 543;
'in front of the house' in Heine & Kuteva, 2002, p. 214)

(4a) na'o-na 'ifi (Kwaio)
'in front of the house' (Keesing, 1991, p. 335.
in Heine & Kuteva, 2002, pp. 141-2)

(4b) na'o-na omea
'before the mortuary feast'

A central mechanism of semantic-pragmatic changes in grammaticalization is that of **''inference''**, or of the ''conventionalization of implicature''. If a grammaticalizing expression routinely occurs in an environment in which a certain inference is made, the expression may gradually become associated with such inference, to the extent that it is incorporated in the meaning of the expression (see Bybee *et al.*, 1994 for references and discussion). This is the case in the development of a causal meaning out of the originally temporal connective 'since' (Traugott & König, 1991, pp.

194-5; Bybee *et al.*, 1994, pp. 196-7) below. In (5a), only temporal meaning is expressed. However, the frequent occurrence of sentences like (5b) offered the contexts allowing a causal inference out of the temporal meaning, to the extent that sentences involving only a causal interpretation of the connective may nowadays be produced (5c).

(5a) I have done quite a bit of writing since we last met.

(5b) Since Susan left him, John has been very miserable.

(5c) Since you are not coming with me, I will have to go alone.

(Bybee *et al.*, 1994, pp. 196-7)

Grammaticalizing expressions also show a shift from encoding some particular semantic content to encoding generalized, reduced semantic content. This semantic change has been widely termed **‘‘semantic bleaching’’** (Givón, 1975), ‘‘desemanticization’’, or ‘‘loss in meaning content’’ (Heine, 2003, pp. 578-9). The construction in (6a), for example, has lost the semantic component of volition, as traced in (6b), and expresses future events in a less constrained fashion, involving inanimate subjects, as in (6c).

(6a) [WILL VP_{inf}]

(6b) I will marry you.

(6c) The bomb will explode.

In the same vein, there is a gradual **increase in generality** in the range of contexts in which grammaticalizing phrases occur (Heine, 2003; Bybee, 2006). In Shakespearean English, for example, the meaning of ‘be going to’ involved a subject literally travelling to a location in order to do something. The subject position could thus be occupied only by a noun phrase denoting an animate, mobile entity, and the following verb was required to be a dynamic verb, as in (7a). Via grammaticalization, a broader

range of appropriate noun and verb phrases were made available for the subject and verb positions, respectively. Thus, non-animate and non-mobile entities may nowadays fill the subject slot, and a broader range of predicates may fill the verb slot (Bybee, 2006), as in (7b).

(7a) I am going to see the king.

(7b) The tree is going to lose its leaves. (Bybee, 2006)

‘‘Chunking’’ (e.g. Boyland, 1996; Bybee 1998), or ‘‘coalescence’’ (Lehmann, 1995[1982]), involves the increase in ‘‘syntagmatic bondedness’’, or ‘‘cohesion’’ (Lehmann, 1995, p. 147-8) of the signs with frequently co-occurring ones in the syntagmatic axis. Chunking may involve ‘‘cliticization’’, i.e. the transition of a sign from a juxtaposed relation with another one to its subordination to the adjacent accent. In (8a), the once independent Latin coordinator ‘que’ turns into an enclitic attached to the first word of the second conjunct in the sentence. Chunking may also involve ‘‘agglutination’’, i.e. the affixation of the sign to another, co-occurring one as a bound grammatical morpheme. In (8b), the Romance synthetic future was grammaticalized from the [infinitive + have] periphrastic construction. Finally, chunking may involve ‘‘fusion’’ or ‘‘merger’’**,** i.e. the complete syntagmatic unification of the sign with an adjacent element (Lehmann, 1995 [1982]). In (8c), a verb of volition, in the course of becoming a future marker from Ancient to Modern Greek, has been merged with its co-occurring subjunctive marker ‘na’ into a single item (e.g. Tsangalidis, 1999).

(8a) Qui ex patre filioque procedit. (Latin)
REL.1SG. from father.SG.ABL son.SG.ABL-and proceed.PRES.1SG
‘Who proceedeth from the Father and the Son’

(8bi) cantare habet (Latin) > cantar ha > cantará (Spanish)
sing.INF.PRES have.3SG sing.fut.3SG
‘He will sing’

(8bii) cantare habeo (Latin) > je chanter-ai (French)
sing.INF.PRES have.1SG PRO.1SG sing.fut.1SG
'I will sing'

(8biii) cantare habeo (Latin) > cantar-ei (Portuguese)
sing.INF.PRES have.1SG sing.fut.1SG
'I will sing'

(Fleischman, 1982, p. 115)

(8c) (ε)θέλω 'ίνα (A. Greek) > θέλω ίνα > θέλω να > θένα > θενα > θα (M. Greek)
(e)t^{h}elō 'hina >'θelo 'ina >'θelo na > 'θena > θena > θa
want.1SG.PRES to
'I want to'

Similarly to chunking, **''phonetic attrition''** (Givón, 1979), ''erosion'', or ''phonetic reduction'', represents a reductive formal change, a ''loss in phonetic substance'' (e.g. Heine, 2003, pp. 578-9; Heine & Kuteva, 2003), this time, inside the sign per se. For instance, the definite direct object marker 'ra' in Farsi is reduced to '-o' after consonants, as in (9) below:

(9) in ce'tαb- rα xæri'dæm > in ce'tαb-o xæri'dæm (Farsi)
DEM book OBJ bought.1SG.
'I bought this book'

Because of changes like the above, linguistic units **''decategorialize''** (Hopper & Traugott, 1993), i.e. they lose morphosyntactic properties characteristic of the class of their source forms/syntactic categories, often including their previously independent word status, and become members of more grammatical categories (Heine & Kuteva, 2002; Heine, 2003, pp. 578-9). For instance, full verbs may become auxiliaries (e.g. 'have', 'do') or prepositions (e.g. 'during'), and demonstrative pronouns may become relative clause markers (e.g. 'that').

3.3. From Grammaticalization to Neuroscience

However, the attempt to investigate the neurolinguistic foundations of such processes meets a lack of interdisciplinary discourse among grammaticalization theory, psycholinguistics and, a fortiori, cognitive neuroscience. The reflections of grammaticalizationists on the psycholinguistic foundations of their own ontology remain quite poor – some bright exceptions are found in the work of Bybee (e.g. Bybee, 2007a), as well as Kempson and Cann (2007), and Rosenbach and Jäger (2008). Psycholinguists, in the same vein, have not dealt much with the manifestation of linguistic phenomena at larger historical time-scales – work by Jurafsky and colleagues (e.g. Jurafsky *et al.*, 2001) is a characteristic exception. Finally, while research on the neurocognitive foundations of psycholinguistic phenomena is certainly much richer, little has been said about the involvement of dynamic neurolinguistic components in historical language change – Givón's work (recently summarized and updated in Givón, 2009) marks the only attempts for such a synthesis.

3.3.1. The Psychological Demonstrandum

While it is a well-acknowledged fact that ''[h]istorical linguists aim to explain language change, ultimately in terms of properties of the human mind'' (Aitchison, 2003, p. 736), psychology remains far from becoming a sister discipline of historical linguistics. Awareness of this lack of discourse is often manifested in the field of grammaticalization theory:

> […]what we are actually talking about is the observable result that is based on our comparison between distinct uses, the historically earlier and the historically later […] What remains uninvestigated is the unobservable aspect of the psychology of language use, i.e. the psycholinguistic mechanism that triggers the process […] (Kuteva, 2001, p. 167f).

Those ‘‘psycholinguistic mechanisms’’ remain, overall, elusive, and pleas have often been made for their investigation (e.g. Bybee *et al.*, 1994; Kuteva, 2001; Tomasello, 2003; Hoefler & Smith, 2008, 2009). In the following sections, a preliminary attempt is presented in the direction of satisfying this desideratum from a neurolinguistic perspective.

3.3.2. The Neurocognitive Demonstrandum

Constructing a neurolinguistics of grammaticalization is tantamount to identifying the causal chain running through the different ontological domains presented in figure 3.1: neurocomputations, psycholinguistic processes, and grammaticalization phenomena. In view of the poor discourse between grammaticalization theory and psycholinguistics-neurolinguistics, the endeavor here needs to follow two basic steps.

The first step in such direction would be the identification of the neural foundations of the domain-general cognitive mechanisms underlying grammaticalization operations. Grammaticalizationists have often made, systematically or not, considerations on the domain-general mechanisms that their ontology reflects (see section 3.5.2 below). Those mechanisms are based on particular neural computations – a basic materialist commitment of modern cognitive neuroscience. One may thus arrive at the neural foundations of such domain-general mechanisms (figure 3.1.A). The second

step, which is toward the ultimate desideratum in this chapter, would be to identify the particular psycholinguistic mechanisms underlying grammaticalization phenomena, and construct a neurolinguistic grounding for them (figure 3.1.B).

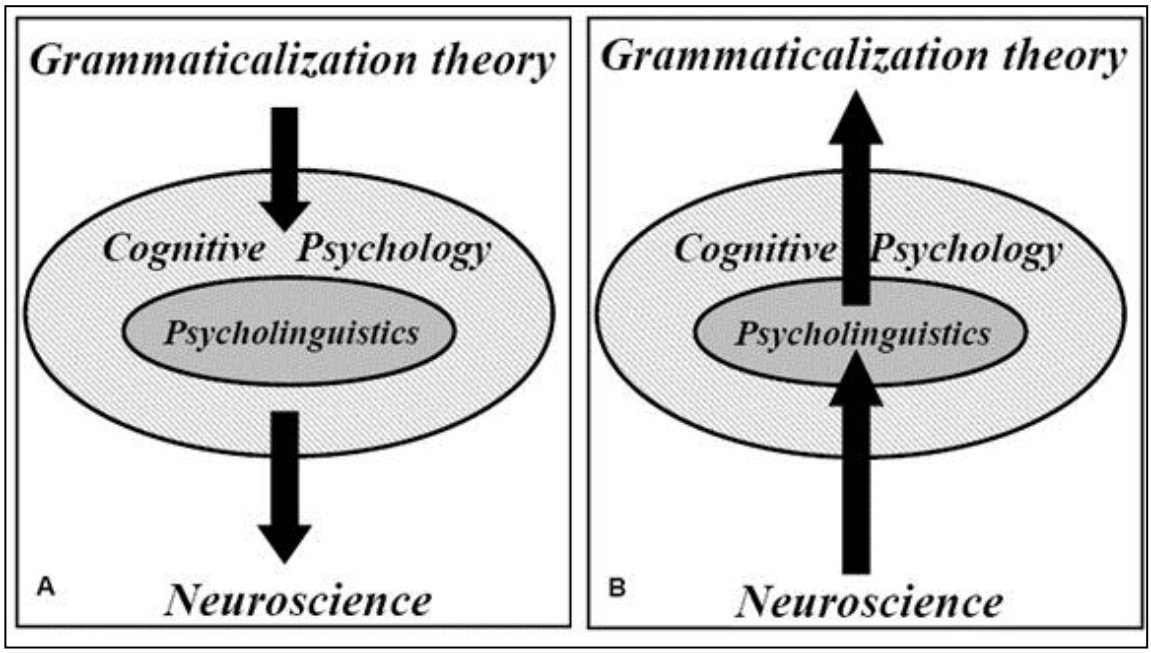


Figure 3.1: The ontological-epistemological chain. A. The first step: The neurocognitive grounds of the psychological changes reflected by grammaticalization operations provide some first insight into the neurocognition of grammaticalization. B. The desideratum: A concrete neurolinguistic model explaining the intra-generational psycholinguistic mechanisms introducing grammaticalization variants.

3.3.3. On Grammaticalization as an Explanandum

However, a number of clarifications are necessary on what such an approach can and cannot explain. To begin with, the account here remains safely far from acquiring a psychological or biological reductionist stance that would ignore the social aspects of grammaticalization or its multigenerational depth. Factors like literacy, for example, might have a direct impact on the material with which the psycholinguistic apparatus interacts, by accelerating or inhibiting the propagation of grammaticalized variants. In

the same vein, a central psychological factor in grammaticalization, such as the cognitive effects of repeated instances of processing (sections 3.4 and 3.5.2 below) would not predict the particular semantic properties of the constructions undergoing grammaticalization:

> [r]epetition alone […] cannot account for the universals of grammaticization. The fact that the same paths of change are followed in unrelated languages has multiple causes. It is not just the fact of repetition that is important but in addition what is repeated that determines the universal paths. The explanation for the content of what is repeated requires reference to the kinds of things human beings talk about and the way they choose to structure their communications (Bybee, 2003, p. 622).

Furthermore, an explanation of such a multigenerational linguistic process should at all times be constrained to the neurolinguistic mechanisms available in the ontogenetic timescale of the interlocutors. This would save the approach from any teleological fallacies. Evidently, the historical time-scale that these processes necessitate to become manifest cannot allow any particular neuroanatomical structure to foretell the psycholinguistic or sociolinguistic phenomena that will occur in subsequent generations.

Most importantly, the psycholinguistic phenomena presented above (section 3.2.2) need not be particular to grammaticalization. Phenomena accompanying the historical emergence of new constructions appear independently, outwith grammaticalization contexts. For instance, items may lose their formal independence (10a-c) or any components of their original semantic representations (10b-c) within the context of lexicalization and idiomatization,[44] with no particular construction undergoing grammaticalization.

[44] Lexicalization and idiomatization refer in general to the historical process by which morphologically complex (words and) constructions tend to behave unpredictably with respect to their compositional

(10a) Baby sitter > babysitter; trouble shooter > troubleshooter

(10b) God be with you > Goodbye!

(10c) qor'bαn-e šo'mα (Farsi)
sacrifice.SGEZ POSS.2PL.
(lit. 'I am your sacrifice')
'thank you/ goodbye'

Such observations have in fact added credibility to the idea that "there is no such thing as grammaticalization, at least in so far as it might be regarded as a distinct grammatical phenomenon requiring a distinct set of principles for explanation" (Newmeyer, 1998, p. 226), or that "[…] grammaticalization is derivative, epiphenomenal, and has no independent status of its own" (Campbell, 2001, p. 151; similar claim in Janda, 2001, p. 266). According to this view, linguists should not ''invite the conclusion that some dynamic is at work in grammaticalization that cannot be understood as a product of [independently occurring semantic/phonetic changes]'' (*ibid*, p. 259; see also Newmeyer, 2001, section 3).[45]

Against these claims, the uniformity of grammaticalization has been defended in grossly two ways, epistemologically, and ontologically. On the one hand, it is argued that grammaticalization theory is a unified field of research in that it involves questions around the emergence of new constructions and the transformation of 'lexical items' into 'grammatical' ones. Grammaticalization theory is thus ''not one single monolithic system […], but describes a range of related approaches and basic issues'' (Haspelmath, 2004, p. 23). On the other hand, in a stronger fashion, it is argued that

properties, and eventually acquire representation as a unit in the mental lexicon. The reader will notice in the following discussion that the same neurocognitive mechanisms may apply to a great extent to idiomatization. For a discussion of the strong relationship between idiomatization and grammaticalization, see Himmelmann (2004).

[45] The issue of the uniformity of grammaticalization often co-occurs with that of its unidirectionality, i.e. whether historical change always proceeds from lexical to grammatical and from less to more grammatical. These issues are beyond the scope of the investigation. Suffice it, though, to mention that evidence of 'degrammaticalization' is multiply less frequent than that for grammaticalization (e.g. Haspelmath, 2004). This has in fact been attributed to the nature of the cognitive mechanisms underlying grammaticalization (e.g. Lehmann, 2004; see section 3.5.2 here).

grammaticalization comprises a uniform ontology, i.e. ‘‘a macro-level phenomenon which cannot be reduced to the properties of the corresponding micro-level phenomena’’ (Haspelmath, 2004, p. 26). Instead, this phenomenon would involve ‘‘strong correlations between phonological, syntactic and semantic-pragmatic changes’’ (*ibid*), suggestive of a common cognitive core (e.g. Lehmann, 2004; see section 3.5.2 below).

Would the neurolinguistic mechanisms discussed below necessitate the ontological uniformity of grammaticalization? Not necessarily. To begin with, these mechanisms are not proposed to monopolize the grounding of grammaticalization. Conceivably, the latter may comprise a heterogeneous range of processes, with different such mechanisms involved to different extents. However, certain conditions of language use may invite processes of the same computational properties. The extent to which these conditions are fundamental in grammaticalization adds equivalent explanatory significance to the neurolinguistic mechanisms they involve. Indeed, the common cognitive core that is often found in the literature to underlie these ‘‘strong correlations’’ will provide the foundation for step 1 of the approach (section 3.3.2). Similarly, this cognitive core and the neurolinguistic mechanisms discussed below need not exclusively ground grammaticalization processes. The extent to which these mechanisms may apply to other historical processes is open for future investigation. The lack of grammaticalization-particular phenomena would thus remain far from suggesting that a neurolinguistics of a historical phenomenon is impossible. Above all, the aim of the endeavor here is not simply to explain those phenomena accompanying grammaticalization (e.g. chunking, phonetic and semantic attrition), which can also be found in other historical changes. The goal is rather to seek the neurolinguistic foundations of the processes by which particular constructions, in certain conditions of language use, gradually deviate from their generic schemata and give rise to new, equally generic ones.

3.4. Grammaticalization is not 'Glossogenetic' [46]

Before continuing with the two basic steps for the neurocognitive grounding of grammaticalization (section 3.3.2 above), it is important to consider what in particular those neurocognitive mechanisms will instantiate. Will they ground processing limitations and learning biases that compromise the transmission fidelity of language inputs that fail to adaptively abide by them? Or will they instantiate computations that adaptively modify the processing of language input in a direct fashion? This does not boil down to an issue of different perspectives on the same ontology. As it shall be shown below, different cognitive mechanisms are involved in each case, triggered in different conditions, with different outcomes in historical language change.

3.4.1. The Glossogenetic Functionalist Framework

Advocating the former, ''glossogenetic functionalist'' principles of language evolution (e.g. Kirby & Hurford, 1997) have been recently suggested to provide a fruitful explanation for grammaticalization changes (e.g. Christiansen & Chater, 2008). Contrary to this position, I will argue that the incorporation of grammaticalization processes in this ontology is far from unproblematic, and that grammaticalization is an explanandum fundamentally incompatible with the glossogenetic functionalist explanans. A briefer version of the argument presented here can be found in Argyropoulos (2010a).

[46] The term 'glossogenetic' is not used here in its broadest sense, by which it refers to the entirety of mechanisms involved in diachronic language change (e.g. Hurford, 2009). Instead, it denotes the principles of adaptation through cultural transmission, which is often termed ''glossogeny'' (Hurford, 1991; Christiansen & Kirby, 2003). It thus refers here to the principles of ''glossogenetic'', as opposed to those of ''phylogenetic functionalism'' in explaining the evolution of language structure (see section 3.4.1).

According to the principles of glossogenetic functionalism, the evolution of linguistic structure is owed to mechanisms by which languages themselves, much like organisms (Christiansen, 1995; Deacon, 1997), are subject to selection pressures. In particular, languages adapt to pressures of transmission in the ‘‘the arena of use’’ (Hurford, 1987), where learning and processing constraints are of catalytic importance: in the absence of sufficient exposure, linguistic input challenging such constraints does not propagate through the transmission ‘‘bottleneck’’ (Hurford, 2002). In this respect, ‘‘it is languages, not language users that are adapting’’ (Kirby *et al.*, 2004, p. 599), since the emergence of more functional forms of communication is ‘‘merely a happy bi-product [sic] of the adaptive mechanism at work’’ (*ibid*, p. 602), which adds weight to the idea that ‘‘language has been shaped by the brain, not the reverse’’ (Christiansen & Chater, 2008, p. 498). Thus, properties of ‘performance’, such as processibility and frequency of exposure gain significant explanatory power for the construction of ‘competence’. A similar idea is found in the ‘‘Performance-Grammar Correspondence Hypothesis’’, according to which ‘‘[g]rammars have conventionalized syntactic structures in proportion to their degree of preference in performance, as evidenced by frequency of use and ease of processing’’ (Hawkins, 2004, p. 3).

Glossogenetic functionalism has thus provided a radical alternative to ‘‘phylogenetic functionalism’’ (e.g. Kirby & Hurford, 1997). The latter explains language evolution on the basis of the phylogeny of the ‘‘Language Acquisition Device’’ (Chomsky, 1965), which adapts to pressures in an environment where successful communication is advantageous. Thus, in the glossogenetic functionalist approach,

> […] the constraints on variation are not assumed to arise directly from the structure of our innate language learning mechanism. Instead, the universals emerge over a historical/cultural timescale from the process of language acquisition and use. This type of explanation relies on the

> principle that language learner [sic] does not necessarily converge on the same grammatical system as the adults in the population. Crucially, the triggering experience that the learner uses will not accurately reflect the linguistic competence of the adults because it is filtered through the ‘‘arena of use’’[…] There are various pressures that operate during communication that will have a selective effect on the different linguistic variants that are being transmitted from generation to generation (Kirby & Hurford, 1997, p. 494).

Accordingly, in computational simulations of cultural transmission, linguistic ‘agents’ form hypotheses on the ‘meaning’ of the data that they produce and observe from other agents. Language, then, changes through misconvergences between the hypothesis that ‘Agent 2’ generates, and the hypothesis based on which ‘Agent 1’ outputs the data that ‘Agent 2’ will eventually observe (e.g. Brighton *et al.*, 2005). In the absence of sufficient training examples, particular form-meaning mappings fail to pass through the bottleneck, and are marginalized or regularized by more learnable and processible mappings. These mechanisms have been used to simulate a growing range of patterns of historical language change, e.g. the emergence of word order universals, where word orders of different processing complexity compete in the arena of use (Kirby, 1999), or the emergence of asymmetric distribution of regular and irregular forms according to frequency, whereby regular forms and irregular ones of high frequency survive in transmission, as opposed to irregular forms of low frequency (Kirby, 2001).

3.4.2. Beyond Formal Misconvergence

I would like now to argue that such mechanisms are problematic when employed for the explanation of grammaticalization operations. In particular, they are susceptible to the same arguments addressed against generativist accounts of grammaticalization (especially Haspelmath, 1998; Lehmann, 2004).

Work on language change explicitly assuming the ‘‘Principles and Parameters’’ (Chomsky, 1981a) framework (e.g. Clark & Roberts, 1993; Niyogi & Berwick, 1997; Briscoe, 2002) has been of paradigmatic value for glossogenetic functionalist models, since it adds weight to the claim that language change is driven by considerations arising ‘‘directly […] from misconvergences […] during language acquisition’’ (Brighton *et al.*, 2005, p. 301). In those studies, such misconvergences occur because the language processor meets severe costs in assigning a particular structure to certain instances of ambiguous input. In view of these difficulties, the processor ‘reanalyzes’ such input according to a considerably more acquirable and/ or processible structure, as in Roberts’ (1993, p. 228-9) ‘‘Least Effort Strategy’’ (see also Lightfoot, 1979, for similar accounts), and a different grammar is thus employed for the acquisition-comprehension of the particular input. More explicitly, Hashimoto and Nakatsuka (2006) assume the position that ‘‘reanalysis’’ and ‘‘analogy’’ underlie grammaticalization (Hopper & Traugott, 1993; Campbell, 2001), and instantiate them in the operations that linguistic agents perform in learning and generalizing their grammar in computational models (Kirby, 2002).

Reanalysis is traditionally defined as a change ‘‘in the structure of an expression or class of expressions that does not involve any immediate or intrinsic modification of its surface manifestation’’ (Langacker, 1977, p. 59). This is accompanied by ‘‘analogy’’, with ‘‘the attraction of extant forms to already existing constructions’’ (Hopper & Traugott, 1993, p. 56). Lexical morphology offers some suggestive examples, as the backformations in (10) below (e.g. Becker, 1993): the morphological structure of

'software', 'hamburger', and 'alcoholic' is reanalyzed as consisting of 'soft-ware', 'ham-burger', 'alco-holic', and, via analogy, new words are created on the basis of this reanalyzed structure. A characteristic case of reanalysis in the syntactic level is shown in (11), whereby a prepositional phrase headed by 'for' (11a) is reanalyzed as belonging to the following infinitival clause (11b) in English (see also Lehmann, 2004, for the same example in German).

(10a) Software: hardware, courseware, netware, wetware,

(10b) Hamburger: cheeseburger, tofuburger,

(10c) Alcoholic: workaholic, chocoholic

(11a) [It would be$_V$ [better$_A$ [for me]$_{PP}$]$_{AP}$ [to slay myself]$_{S\text{-}INF}$ [than to be violated thus]$_{S\text{-}THAN}$]$_S$

(11b) [It would be$_V$ better$_{AP}$ [for me to slay myself]$_{S\text{-}INF}$ [than to be violated thus]$_{S\text{-}THAN}$]$_S$.

(Haspelmath, 1998, p. 324-5; originally from Harris & Campbell, 1995, p. 62)

The insufficiency of reanalysis to explain grammaticalization has been discussed elsewhere (especially Haspelmath, 1998; Lehmann, 2004). In what follows, I briefly provide a number of highlights, and emphasize only those that are central for the argument here. To begin with, grammaticalization has adult language users as the primary agents of change, and not language-acquiring infants (see Croft, 2000, for arguments and references). The contribution of the latter has been instead regarded as catalytic in reanalysis and analogical levelling (Hooper, 1976; Haspelmath, 1998).[47] Independently of the agent carrying out the change, reanalysis and grammaticalization have been shown to often act independently from each other; reanalysis also involves contracting new syntactic relations with sentence elements not related before, whereas grammaticalization does not involve any real change in constituent structure (Haspelmath, 1998; Lehmann, 2004). Furthermore, putative examples of reanalysis in

[47] Schematization mechanisms precisely reflect the analogical nature of such computations (sections 2.3.4.3 and 3.6.4).

grammaticalization are often shown to occur as a relabeling of nodes in constituent structure trees. In the Swahili example in (12b) below, the perfect prefix 'me' originates by the grammaticalization of a verbal head 'meele' (12a) into a verbal affix- this development has been described as a product of reanalysis (Heine & Reh, 1984, p. 102). However, those reanalyses are theory-dependent and epiphenomenal, since they are immediately done away with when described in other syntactic theories (e.g. dependency grammar). Instead, these examples are best described in terms of a gradual increase in bondedness of internal dependencies inside the construction[48] (Haspelmath, 1998).

(12a) $[[\text{mtoto}]_{NP}\ [\text{a-meele}_{\ v}\ [\text{ku-ja}]_{NP}]_{VP}]_{S}$ (Heine & Reh, 1984, p. 102; Haspelmath, 1998, p. 334)
child 3SG-finish come.INF
'The child finished coming'

(12b) $[[\text{mtoto}]_{NP}\ [\text{a-me-kuja}_{V}]_{VP}]_{S}$
child 3SG-PERF-come
'The child has come'

For the purposes of the present investigation, the key differences to note are that, unlike grammaticalization, reanalysis presupposes structural ambiguities in the input, and is an abrupt, categorical phenomenon, which can only operate on already available categories (Haspelmath, 1998). On the contrary, grammaticalization does not result from 'misprocessing'-'misacquiring' the input, and it is a gradual process that may lead to structures and categories that have not existed: ''[…] reanalysis cannot create anything genuinely new. Grammaticalization is capable of exactly this'' (Lehmann, 2004, p. 11). A typical example of the latter would be the development of definite articles in Romance languages out of the Latin demonstrative, as in (13) below. A reanalysis account would fall short in identifying an already available structural analysis whereby the new 'DET(erminer)' category emerges.

[48] See section 3.6.2 for an example, where the increase in such ''bondedness'' arguably translates into an increase in transitional probabilities in processing.

(13a) lupus ille (Latin) > lupul (Romanian)
wolf. DEM wolf.DET
'that wolf' 'the wolf'

(13b) ille canis (Latin) > le chien (French)
DEM. dog DETdog
'that dog' 'the dog'

Similarly to reanalysis, many phenomena of grammaticalization are often seen as cases of structural ''reinterpretation'' (Heine & Kuteva, 2002). For example, the development of complementizers out of demonstratives as in (14) below, has been described as an instance of

> […] reinterpretation of certain patterns of direct speech (e.g. She said that: there is no money) as a main clause + complement clause combination (She said that there is no money), where the demonstrative object of the matrix clause, referring cataphorically to the next clause, is reinterpreted as a marker introducing a complement clause (*ibid*, p. 106).

As above, explaining this transformation as an instance of misinterpretation silently assumes that the representation of a 'marker + complement clause' construction is already available in the interlocutors' grammatical representations.

(14a) Eg sigi tadh: hann kemur (Lockwood, 1968, pp. 222-3)
I say that: he comes
'I say this: he comes.'

(14b) Eg sigi at hann kemur.
I say that he comes
'I say that he comes.'

Instead of addressing the very sufficiency of glossogenetic functionalist mechanisms to explain grammaticalization, criticism so far has rather emphasized the need to simply abandon vertical, language acquisition-based models of transmission for the simulation of historical change (Croft, 2004). However, this premise has been identified as only a methodologically justified simplification (Smith *et al.*, 2003). The problem thus persists, since the 'reanalytic' mechanism of misconvergences in transmission, which is explicitly considered irrelevant in grammaticalization (Haspelmath, 1998), may still be employed in a horizontal transmission model. Change would still be introduced with agents misattributing an underlying syntactic-semantic representation to linguistic strings because of processing and learning limitations that the input strings exceed. The horizontal transmission model in Batali (2002, p. 115) would offer a characteristic example: the sender transmits a signal S to express a meaning M1, following a particular analysis A1 of the mapping from the structure of M1 to the signal S. However, the receiver derives an interpretation M2 for the signal S according to analysis A2 of the mapping from the signal structure to M2. Language change is thus initiated by a discrepancy between A1 and A2, where A2 is more optimal than A1 with respect to the processing capacities of the language agent. This is the case in reanalysis and analogical levelling – not in grammaticalization, unless one adheres to the heavily criticized generativist models like the above.

3.4.3. Beyond Semantic Misconvergence

It has similarly been suggested that representational misconvergences between interlocutors in semantic aspects of discourse provide a trigger for grammaticalization. In some recent proposals, for example, the metaphorical extension often involved in grammaticalization (e.g. from physical to temporal movement in the 'be going to' construction) has been discussed as the production-equivalent of a process of a pragmatic ''reanalysis'' from the hearer's side (Hoefler & Smith, 2008, 2009). However, mechanisms like metaphorical extension are far from intrinsic to grammaticalization

processes, as they pertain to a radically broader range of lexical items, only a minority of which eventually acquires grammatical functions (as example 3 above). For instance, the metaphorical extension of body-part terms to express spatial concepts is arguably not an instance of grammaticalization. In (15) below, for example, some English body-part terms are used to express spatial relations. However, none of these expressions is currently under grammaticalization (Bybee *et al.*, 1994). Metaphorical extension and semantic bleaching rather allow expressions to increase in frequency of occurrence and acquire basic discourse status so as to undergo the formal and semantic grammaticalization changes. However, they do not suffice to explain the core transition from 'lexical' to 'grammatical' that items in grammaticalizing expressions undergo (Haspelmath, 1999, p. 1062). Other researchers similarly find ''no case of generalization changing a lexical to a grammatical expression […], suggesting that generalization as a mechanism of semantic change is not itself the mechanism by which grammaticalization is activated" (Nicolle, 1998, p. 22-23). Furthermore, much like lexical change, and much unlike the conventionalization of new inferences (section 3.2.2), the metaphorical shift from the source (e.g. physical progression) to the target domain (e.g. temporal progression) is rather innovative and abrupt in nature (Bybee *et al.*, 1994, p. 197) – this would explain the appeal of metaphorical extension to reanalysis-based accounts discussed above.[49]

(15) the foot of the bed, the face of the cliff, the face of the clock (Bybee *et al.*, 1994, p. 284).

[49] In contrast to metaphor, which innovatively introduces a novel form-meaning mapping (a 'competence' representation), 'metonymy', or 'the conventionalization of inference', will be seen as a shortcut behavior in discourse processing, and thus as a 'performance change' (section 3.6.3.2.2).

3.4.4. Ontological Dissociation

In delimiting the ontology of glossogenetic functionalist and grammaticalization mechanisms, it is valuable to consider the distinction between the ‘‘conserving’’ and the ‘‘reducing effects’’ of repetition in historical language change (Bybee & Thompson, 2000), respectively. The glossogenetic functionalist approach to language evolution is best construable as objectifying the changes introduced by the conserving effects of repetition.[50] Unlike high-frequency and easy-to-process form-function mappings, lower-frequency and less processible mappings are changed on the basis of analogy with more productive ones (Bybee & Thompson, 2000; Bybee, 2007a). For instance, high-frequency verbs, as in (16a), retain their irregular past-tense morphology, unlike lower-frequency verbs, as in (16b), that are assimilated into a stronger morphological schema (‘-ed’; Bybee, 1985). On the same grounds, English pronouns show a conservative syntactic behavior, and maintain distinct forms for nominative vs dative and accusative case marking, while such case distinctions have become extinct from nouns (Givón, 1979).

(16a) keep/kept, sleep/slept
(16b) weep/wept-weeped, creep/crept-creeped

Suggestively, in their brief discussion of grammaticalization processes, Christiansen and Chater (2008) use Hare and Elman’s (1995) simulation as an example of historical language change explained by learning and processing constraints: In Hare and Elman (1995), a connectionist network modelled the historical development of the English verb inflection from its past-tense system in Old English up until the modern times. A network was taught a data set of the Old English verb classes, but learning was stopped before reaching asymptote, and the output of each such network was used as the teacher of a new net. In that way, vertical language transmission was modelled. As a result, the errors in the first network were passed on to become part of the data set of the

[50] Haspelmath’s (2004) principle of ‘‘survival of the frequent’’ offers similar insights.

second. Certain input-output mappings were easier to acquire, such as the highly frequent patterns, and those sharing phonological regularities. The patterns that were hardest to learn led to the most errors, and over time were ‘‘regularized’’ to fit the more regular and more productive patterns. The phenomenon thus described is an example case of analogical levelling and/or conservation processes in the acquisition of verbal morphology. As such, it is far from relevant with grammaticalization processes to encourage a similar approach to the latter, as Christiansen and Chater (2008) imply.

Grammaticalization changes, on the other hand, are products of the multi-level ‘‘reducing effect’’ of repetition (Bybee & Thompson, 2000).[51] Automatization of performance in language processing represents the most widely supported cognitive core of grammaticalization (see section 3.5.2 below). It reflects domain-general, non-species-specific adaptive responses of the brain to repeated behavioral repertoires generated by particular neurocognitive mechanisms (see sections 3.5.1, and 3.5.3 below). Here, linguistic change does not amount to misprocessing of ambiguous input, but is the result of the user’s adaptive minimization of cognitive and sensorimotor costs in language processing. The dissociation of those two kinds of change is summarized in figure 3.2. This is illustrated by Hooper’s (1976) observation that low-frequency items are affected by analogical levelling first (16 above), whereas sound change in the form of phonetic attrition affects high-frequency items first. For instance, deletion of final /t/ and /d/ in American English is (at least partly) determined by frequency (Bybee, 2002d), with high-frequency words (17a) being more susceptible to this change than low-frequency ones (17b).

(17a) just, child, grand

(17b) conformist, cuckold, blend

[51] The two processes are not directly mapped to the ‘conserving/reducing effects’ divide, but only to the changes involved in each case: in grammaticalization, language changes as a product of multi-level reduction, whereas, in the glossogenetic functionalist framework, language changes as a result of unsuccessful conservation: in the latter, mechanisms of analogy, reanalysis, and generalization kick in. This is not incompatible then with the fact that grammaticalization requires and richly interacts with the conserving effects of repetition (e.g. section 3.6.4 below).

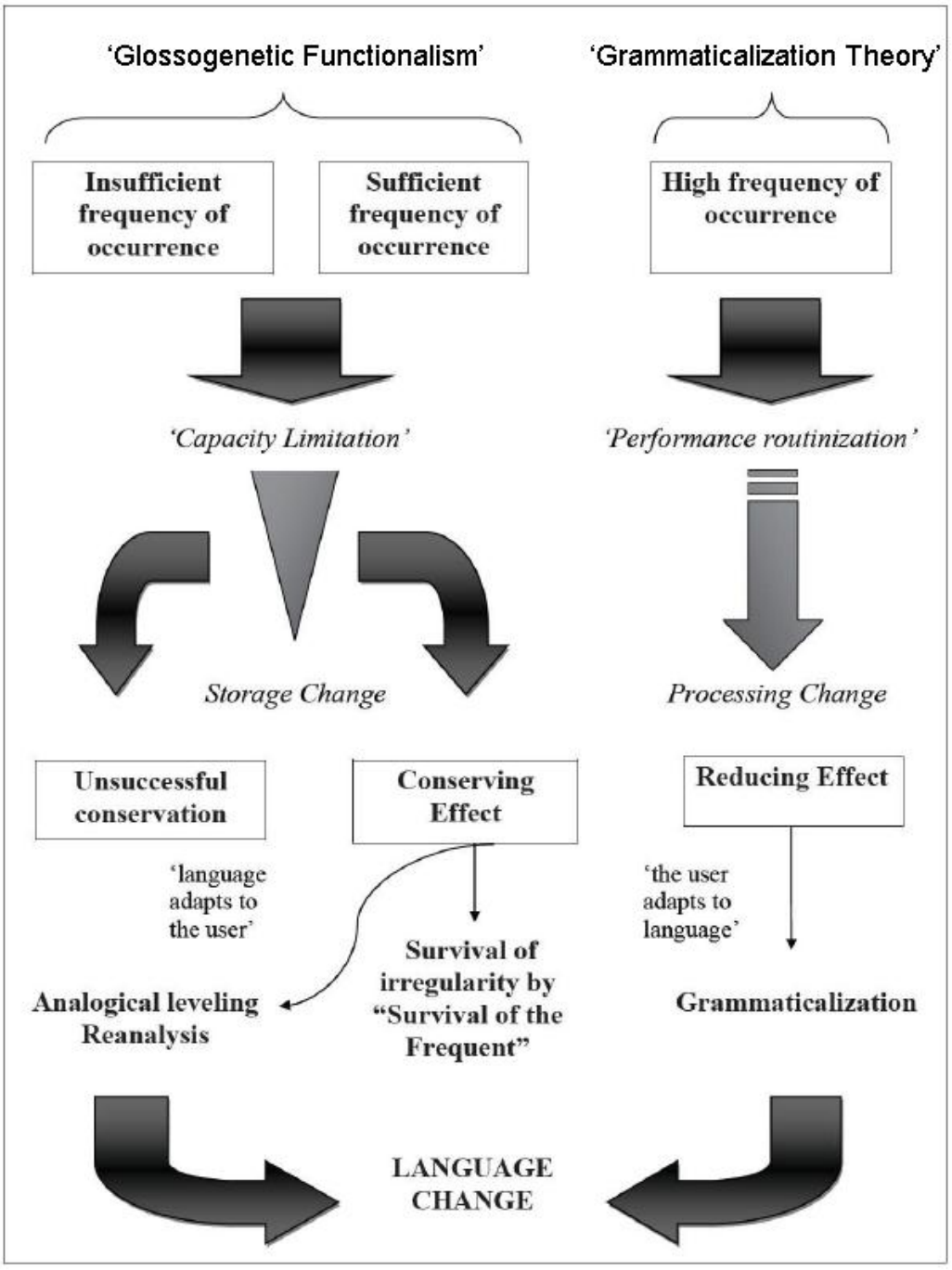


Figure 3.2: Dissociation between 'glossogenetic functionalist' and 'grammaticalization theory' mechanisms, on the basis of the distinction between unsuccessful conservation and successful reduction. The 'sufficient vs high frequency of occurrence' distinction here is in line with that made in Bybee (2006): constructions of ''high frequency'' are simply conventionalized, whereas constructions of ''extremely high frequency'' undergo grammaticalization (see section 3.5.2 below). Figure adapted from Argyropoulos (2010a, p. 7) with permission © 2010 World Scientific.

Thus, despite their emphasis on 'performance' properties[52] (e.g. frequency, processing difficulty of items), glossogenetic functionalist approaches crucially remain ''competence-based'', much like the generativist explanation of grammaticalization: strings of input are re-assigned and/ or are produced according to a different structural representation. In other words, language change in the glossogenetic functionalist framework is necessarily seen as the product of competence changes owing to the performance properties of the transmitted input. On the contrary, ''performance changes'', without the precedent of a corresponding ''competence change'', are beyond the scope of the glossogenetic functionalist approach. However, grammaticalization relies precisely on ''performance changes'' (Haspelmath, 1998), that do not result from altered competence representations. Changes in the latter may only follow after the coding of the introduced variant by a new cohort of speakers (see section 3.6.4 below).

In closing this section, then, the question in the beginning can now be answered: the nature of grammaticalization is such that its underlying neurolinguistic components should account for direct, optimizing changes in processing – not for the learning and/or processing biases and limitations by which misconvergences in transmission occur. The adaptive responses of the brain to reoccurring instances of behavior, described in the literature as 'automatization' or 'routinization', will be examined in the next section as the primary cognitive mechanism underlying performance optimization and grammaticalization.

[52] The 'competence-performance' distinction, is, of course, far from accepted in the functionalist camp, and, a more appealing distinction could be made on the neutral grounds of ''cognitive representations and usage'' (J. Bybee, personal communication, 5 December 2007). However, Haspelmath's (1998) ''competence/performance change'' distinction is adopted here, in a similar attempt to establish common grounds for discussion among Chomskian, 'glossogenetic functionalist', and functionalist accounts of grammaticalization.

3.5. Grammaticalization, Automatization, and the Cerebellum

In constructing the neurolinguistic foundations of grammaticalization, the first step, as proposed earlier (section 3.3.2), is to examine the cognitive mechanisms underlying the historical phenomenon, and search for their candidate neurocognitive foundations. Automatization has been in the spotlight of cognitive considerations on grammaticalization, and offers a characteristic case of a performance change (section 3.4.4). After defining automatization, I briefly present such considerations on automatization and grammaticalization, and then discuss some highlights on the significance of the CB for automatization. This will provide the grounds to closely examine the significance of the NCBKFLP for grammaticalization from a neurolinguistic perspective.

3.5.1. Automatization

Automatization involves the practice-induced minimization of cognitive-attentional and mechanical-executional costs in task-performance. Similarly, an 'automatic process' is traditionally defined in terms of the activation of a sequence of nodes that ''nearly always becomes active in response to a particular input configuration'', and that ''is activated automatically without the necessity for active control or attention by the subject''. In general, automatic processes ''operate through a relatively permanent set of associative connections […] and require an appreciable amount of consistent training to develop fully'' (Schneider & Shiffrin, 1977, p. 2). An automatized modality has often been contrasted with an attended processing one, which is manifest in the acquisition of new skills. A fully automatized skill does not require conscious effortful monitoring or guidance, and shows little or no decrement in the presence of other demands upon conscious processing. Automatic process thus allow for

the parallel execution of a (usually less automatized) task, such as having a conversation while skillfully typing a text (Rumelhart & Norman, 1982).

By contrast, a cognitively controlled procedure requires increased levels of attention and guidance of performance. This is demonstrable in the initial phase of skill acquisition and novice performance: a new pianist, at every single key they strike, they have to rely on attentional-perceptual cues from the environment for the next key they should press. Along with such capacity limitation, however, controlled processes have the significant advantage of allowing for flexible alteration and application of a behavior in novel situations, much unlike the context-particular automatic repertoires (Schneider & Shiffrin, 1977, p. 2-3).

3.5.2. Grammaticalization as Automatization

As already suggested, grammaticalization processes involve a multi-level reduction (section 3.4.4), in other words, ‘‘an evolution whereby linguistic units lose in semantic complexity, pragmatic significance, syntactic freedom, and phonetic substance’’ (Heine & Reh, 1984, p. 15). A conditioning factor for those reductions, also discussed above, is the high frequency of repetition of the grammaticalizing construction: ‘‘[i]n grammaticization, changes in phonology, semantics, and structure occur in extremely high-frequency constructions’’ (Bybee, 2006, p. 730).

This property has provided the grounds for the definition of grammaticalization as ‘‘the process of automatization of frequently occurring sequences of linguistic elements’’ (Bybee, 2007b, p. 969). Indeed, these mechanisms, termed ‘‘automatization’’, ‘‘routinization’’, ‘‘ritualization’’, ‘‘habituation’’, comprise the most widely held cognitive core of grammaticalization processes (e.g. Givón, 1979; Haiman, 1994; Bybee, 1998; Haspelmath, 1999; Lehmann, 2004). For instance, several

similarities have been noted between habituation in organisms and ritualization in social practice on the one hand, and grammaticalization processes on the other. 'Habituation', i.e. the process by which an organism ceases to respond to the same extent to repeated stimuli, and 'ritualization', i.e. the process by which a cultural artifact loses its original force and significance, have been discussed as the behavioral-cultural foundations of semantic bleaching (Haiman, 1994; section 3.2.2 here). In the same vein, formal changes occurring in grammaticalization, such as phonetic attrition and chunking (section 3.2.2), closely resemble changes in the ontogenetic course of skill automatization (Boyland, 1996). With the repetition of particular behavioral sequences, formerly separate gestures are integrated into an automated unit, and forms are reduced through the weakening of the individual gestures comprising the act (e.g. Haiman, 1994; Bybee, 1998). Similarly, in Lehmann's (1995[1982]) multi-parametric, semiotic approach, grammaticalization is defined as the minimization of the independence of the linguistic sign, the cognitive equivalent of which is the loss of the freedom to manipulate it (Lehmann, 2004). Automatization, then, stands as a ''[…] psychological correlate to grammaticalization as a structural notion'' (*ibid*, p. 25). These automatization processes then arguably comprise an explanatorily promising cognitive core of grammaticalization. Their neural grounding would thus provide the first insights into the neurolinguistic foundations of grammaticalization.

3.5.3. Cerebellar Foundations of Automatization

Despite the findings on the neural structures responsible for the genre of automatization mechanisms, their very connection with linguistic processes, from either a synchronic or a diachronic perspective, remains a poorly addressed desideratum. Suggestively, in the context of grammaticalization theory, chunking-automatization has been seen as a ''physiologically mysterious process'' (Haiman, 1994, p. 10), and, similarly, it has been suggested that ''[…] the physiology of ritualization in human beings is unknown'' (*ibid*, p. 25).

From a cognitive neuroscience standpoint, such statements have been far from the truth for quite a few decades, and can only indicate the lack of interdisciplinary discourse.[53] I thus briefly turn to the description of the widely held view that the CB provides a strong candidate locus for automatization operations. In chapter 2, some basic information was presented on the role of the CB in feedback-free, accurate, and timely motor control, especially in the context of the CB instantiating internal models (section 2.2). Such functions are of outmost importance in the automatization of behavioral tasks, and have explicitly promoted automatization as an instantiation of the ‘‘universal cerebellar transform’’ (Schmahmann & Caplan, 2006).

To begin with, the cognitively controlled and automatic processing circuits have been distinguished on the basis of imaging evidence: the former mainly relies on prefrontal regions, as well as on activity in the right premotor cortex and bilateral supplementary motor area and left CB, while the latter relies on motor circuits, which crucially involve the right CB, along with the supplementary motor area, the left sensorimotor cortex, the left thalamus and the basal ganglia (Lewis & Miall, 2003). In the automatization of a motor task, the motor loop becomes involved, disentangling the prefrontal cortex to employ attentional resources for another task (Jueptner & Weiller, 1998).

On its own, the CB has been for a long time considered a candidate foundation for behavioral automatization. The acquisition of complex, skilled movements has been described to begin as a conscious act almost exclusively under cerebral control, and to involve, via learning, CB circuitry, that takes control of the task (Brindley, 1964). In other words, the CB

[53] The work of Givón (1979-2009) has been the brightest and perhaps only exception in the literature to consistently discuss grammaticalization in the light of the neurobiology of automatized processing. See section 3.8 for further discussion.

> [...] recognizes the contexts in which each "piece" of consciously initiated movement occurs. After repeated tries, it links that context within itself to the movement generators so that the occurrence of the context automatically triggers the movement. Thus, with time and practice, the cerebellum largely controls the process, with little or no help from the cerebrum. The cerebrum and the conscious mind are free to do and think about other things. Control of the task has been shifted from a conscious cerebral cortical process to a subconscious one mostly under the control of the cerebellum (Thach, 1997, p. 600).

Neuroimaging and clinical evidence supports the significance of the CB in shifting pre-training movement performance to a more skilled and automatic stage. This is often captured in the context of dual-task paradigms, where one of the two tasks involved is automatized, and attentional resources for conscious monitoring are allocated to the other. In such paradigms, CB patients are able to improve a learned movement with practice only to a very limited extent, and cannot perform the task in a more automatic stage. Moreover, when attention is focused away from one task, their performance deteriorates to pre-practice levels. Normal controls, on the contrary, show an expected insensitivity to interferences in dual tasks after practice (Lang & Bastian, 2002). The same automatization deficit has been observed in CB disruptions in working memory and word-generation tasks (section 2.4.2 above).

In imaging studies, strong CB activations have been reported in different stages of automatization. In some studies, CB activity is greatest early in motor learning, supporting the idea that the CB is exclusively involved in shifting performance to a more automatic state (Jenkins *et al.*, 1994; Doyon *et al.*, 1996; Jueptner *et al.*, 1997; Toni *et al.*, 1998). In other studies, CB activity is found to increase only when a practiced movement is performed at a later stage (Shadmehr & Holcomb, 1997). Such findings

would entertain the possibility that the CB is directly involved in the execution of a more automatic movement after its acquisition, with early performance of a novel motor task relying more on frontal cortex. The two roles of course are not mutually exclusive and such a dual role for the CB is tenable (see Lang & Bastian, 2002).

As suggested above, a sign of automatized learned movements is their very performance in units comprising of composite synergies of individual moves (section 3.5.2). A characteristic clinical indication of CB deficit, noted very early in the literature, is the ‘‘decomposition of movement’’ (Holmes, 1939), whereby compound movements are broken down into individual moves, and executed in a seriatim fashion. Strikingly, focal cortical lesions give rise to focal motor dysfunctions, pertaining to one or a few joints and muscles, whereas CB lesions spare single-joint movements, but induce such incoordination in multijoint movement (see Thach, 2007, for references). This has given grounds to the proposal that the CB combines single-joint movements into composite ones (e.g. Bastian *et al.*, 1996; Goodkin & Thach, 2003). A closely related proposal has been that the cerebrum specializes in ‘‘propositional’’, consciously planned and initiated movements, and the CB in ‘‘automatic’’, compound movements, triggered immediately in response to particular experiential contexts (Thach *et al.*, 1992; Thach, 1996a; Thach, 2007).[54]

Moreover, within the framework of the ‘‘dyslexic automatization deficit’’ hypothesis (Nicolson & Fawcett, 1990) and the suggestively concomitant ‘‘cerebellar deficit hypothesis’’ (Nicolson *et al.*, 2001), behavioral evidence has been provided for abnormal CB function in a high percentage of dyslexic children. Classic indices of this have been deficits in skill automatization, time estimation, balance, tone, and eye-blink conditioning (e.g. Fawcett *et al.*, 1996; Fawcett & Nicolson, 1999). Within a range of

[54] The CB mechanism underlying this combining operation is debatable (see the Paulin (1996)-Thach (1996b) debate). Some support the idea that parallel fibers compound by trial and error (LTD) particular experiential contexts to the responses, and single response elements into composite units (see Thach, 2007). CB state estimation, on the other hand, in low Kalman gain conditions, would minimize reliance on external feedback, and would thus allow the execution of compound movements (section 2.2.2).

dual-task paradigms, significant difficulties have been shown selectively for dyslexic children in dual-, as opposed to single-task conditions (Nicolson & Fawcett, 1990). Furthermore, in the acquisition and execution of sensorimotor sequences, dyslexics appear to activate the CB less than controls, instead relying significantly more on frontal lobe functions (Nicolson *et al.*, 1999). Recent studies also provide neuroanatomical evidence indicative of CB abnormalities in dyslexic brains. In particular, significant differences have been found in Purkinje cell size in medial areas of the posterior CB lobe, and also in the Purkinje cell size distribution in the anterior and posterior CB lobe and the cells in the inferior olive (Finch *et al.*, 2002).

This section thus completed the first step of the approach (section 3.3.2). Grammaticalization changes are widely held to rely on automatization processes in language processing at the intra-generational level of language transmission (section 3.5.2). CB computations were shown to be a major component of neurocognitive automatization processes (section 3.5.3). Given the cytoarchitectural homogeneity of the CB (section 1.2.3) and its reciprocal projections to language-related cortical loci (section 1.3.2), CB computations may arguably have the same contribution to language processes. The striking lack of practice-induced facilitation that NCB patients show in word-generation task-performance (see section 2.4.2) already stands out as a meeting point for the three disciplines here, namely, NCB computations, language processing, and grammaticalization operations. Moreover, the fact that the CB has a metasystemic role in a wide range of behavioral domains (section 1.4), and the fact that CB lesions make task performance only suboptimal (section 2.4.4.4), promote the CB as a 'competence-peripheral', 'performance-optimizing' machine. That was a necessary premise in identifying a suitable neuroanatomical foundation for grammaticalization operations (section 3.4.4). The discussion so far thus provides the grounds for investigating the hypothesis below, namely, that the NCBKFLP has a fundamental role in introducing formal and semantic grammaticalization changes as a product of routinized language processing.

3.6. Neocerebellar Neurolinguistics of Grammaticalization

The next section reflects a preliminary attempt to take the second step in constructing a neurolinguistics of grammaticalization (see section 3.3.2), on the basis of the preceding discussion. In what follows, I discuss how grammaticalization phenomena involve psycholinguistic mechanisms that arguably employ the NCBKFLP. A common thread running through the argument below will be that a shift occurs from categorical towards associative computations in routinized multi-level language processing. The instantiation of associative processes, as well as the control of the trade-offs between the two, were argued in the previous chapter to be the core functions of the NCBKFLP (section 2.3.4.3).

3.6.1. Neocerebellar Memory of Dialogue Dynamics

The origins of these multi-level changes described in the beginning of the chapter (section 3.2.2) have been held, especially in the functionalist camp, to occur at the intra-generational level of routinized adult dialogical interactions (sections 3.4 and 3.5.2). Dialogue involves, at first blush, a very demanding task. For example, interlocutors are required to both comprehend and produce in turns or even in tandem, to respond rapidly, and to comprehend elliptical utterances, often under noisy environments. However, there are psycholinguistic mechanisms that make almost everyone capable of holding a conversation (Garrod & Pickering, 2004). Due to the mediation of covert imitation mechanisms (Pickering & Garrod, 2007) and the unconscious multi-level priming of representations (Pickering & Garrod, 2004), interlocutors quickly develop similar activation patterns for sounds, words, grammar, and interpretations – i.e. their representations become aligned. Words and constructions are repeated, and, eventually, routinization phenomena occur, whereby interlocutors

converge on consistent interpretations of forms with great speed, often without explicit negotiation (e.g. Garrod & Anderson, 1987; Garrod & Doherty, 1994).

Dialogue may thus provide an empirically tangible process to capture the more macroscopic grammaticalization phenomena. This has been recently entertained in grammaticalization literature: dialogical routinization has been considered to provide ‘‘the microstructure of linguistic change’’ (Rosenbach & Jäger, 2008), and ‘‘a formal reconstruction of grammaticalization’’ (Kempson & Cann, 2007, p. 76). Repetitious contexts increase predictability of the items perceived-comprehended and articulated-produced.[55] Particular instances of constructions are repeated within the same and across different dialogical contexts, in the same form, inviting the same inferences. This sets up repeated, deterministic input-output relations in processing, e.g. between acoustic stimuli, word strings, and lexical semantics on the one hand, and phonological representations, syntactic structures, and utterance interpretations on the other:

> […] utterances involve an increasing proportion of expressions whose form and interpretation is partly or completely frozen for the purposes of the conversation […] Such routinized expressions are similar to stock phrases and idioms […], except that they only ‘live’ for the particular interaction (Garrod & Pickering, 2004, p. 10).

Thus, if historical phenomena like idiomatization and grammaticalization rely on casual adult language use, and if dialogue represents the fundamental context of language use, then it is only conceivable that certain routinization effects live well beyond ‘‘the particular interaction’’ (*ibid*). Word sequences, structural assignments, and

[55] Note that in Pickering and Garrod (2007), emulative processes are predominantly discussed in the context of their contributions to interactive alignment and in establishing a ‘‘shared reality’’, by representing one’s interlocutor’s mental states (Pickering & Garrod, 2009a). This aspect of imitative language emulation is beyond the scope of the present investigation (section 2.5); on the contrary, the emphasis here is on the computations of internal models after such repetitious contexts are established.

interpretations occurring frequently across dialogical contexts would be the strongest candidates for ultimate storage and conventionalization (see section 3.6.3.1.2 below).

3.6.2. The Neocerebellar Associative Shift

As explained in the previous chapter, NCB internal models may acquire the dynamics of the linguistic processes where particular inputs frequently co-occur with particular outputs (sections 2.3.3 and 2.3.4). In dialogical contexts, the aforementioned mechanisms of repetition priming, alignment and covert imitation give rise to such highly repeated formal and semantic-pragmatic processing sequences. NCB internal models may copy the input-output dynamics of unification processes of repeatedly processed constructions. Reference to categorical properties of these constructions, i.e. to their schematic constructions, is thus attenuated: instead of searching the 'construct-i-con' instantiated in the Memory component, the controller-Broca's area may now perform a constrained search on the compact NCB associative look-up table to perform its unifications (section 2.3.4.3).

In Argyropoulos (2008b, June), I had argued that the very dynamic construal of ''pseudosyntax'' as the process stage of ''syntactic habits'' (Townsend & Bever, 2001) presupposes the involvement of adaptive 'demotions' from stage 2 to stage 1 operations that routinization of language processing would induce. In LAST, highly inflected languages employ such rich grammatical morphology as ''surface cues'' to complete the first stage of sentence processing, which works toward the minimization of the search space for the retrieval of a particular form-meaning mapping (Townsend & Bever, 2001).[56] On the other hand, grammatical morphology is the product of

[56] Once again, it should be emphasized that the hypothesis is not inherently committed to a two-stage model of sentence comprehension. Instead, properties of the first stage in such models are used here to illustrate those of a low Kalman gain modality of processing (see section 2.4.4.1).

grammaticalization processes, and grammaticalization processes are triggered in intra-generational dialogical contexts (sections 3.4.4, 3.5.2, and 3.6.1). Thus, in order to explain how lexical items have been transformed into such surface cues, a necessary trade-off is required to be at work between those two stages, which may lead eventually to an adaptive demotion. These first-stage processes were argued to involve NCB computations (section 2.4.4). The trade-off, in the same vein, is a function of the adaptive combination of categorical, cortico-cortical, and associative, cortico-cerebellar computations that the NCBKFLP performs (section 2.3.4.3). Suggestive evidence for the possible involvement of NCB circuitry in constrained search in language processing was also presented in the previous chapter (section 2.4.4.2).

A similar transformation is assumed in Deacon's (1997) cognitive-semiotic framework. Historical language change is discussed as having stranded certain symbolic items to indexical, subcortically instantiated ''cues'';[57] as described in the previous chapter, the latter help automatize language processing by minimizing the search space for the application of the highly distributed, ''symbolic'' grammatical rules (see section 2.4.4.1 above):

> [i]n the interest of rapidly and efficiently organizing such highly distributed associative processes, some symbols had to be stripped of all but the vaguest symbolic in order to provide a set of automatic "switches" for shuttling symbolic work to the appropriate regions of the brain. (Deacon, 1997, p. 299).

Interestingly, constrained search mechanisms have been explicitly proposed to underlie language processing changes initiating grammaticalization operations. In a more concrete, computationally-based and psycholinguistically-oriented parsing

[57] In section 3.6.3.2.2 below, I also discuss how indexical semiotics strikingly reflects NCB associative computations.

framework, Kempson and Cann (2007) assume a distinction between ‘‘a generally available tree-growth process’’ (*ibid*, p. 76) and a ‘‘lexical look-up mechanism’’ (*ibid*, p. 95). A shift from the former to the latter is used to account for the procliticization of object pronouns in the transition from Medieval to Renaissance Spanish, as in the example (18) below. This shift is ‘‘[…] a natural subsequent step of routinization’’, involving the calling up of ‘‘[…] actions associated with the verb together with those associated with the clitic with a single lexical look-up mechanism’’ (Kempson & Cann, 2007, p. 95).

(18) mas los rompan luego (in Kempson & Cann, 2007;
but CL break.3PL afterwards from Bouzouita, 2002).

‘But break them afterwards.’

The distinction between those two computations, and the shift in processing modalities proposed are directly equivalent to those between the categorical, cortico-cortical, and the associative, cortico-cerebellar computations in language processing, along with the reduction of the NCBKFLP gain proposed here. The conditions in which this shift is triggered are also the same, i.e. routinization, repeated exposure to the same input-output parsing sequence.

3.6.3. Multi-Level Reduction

In processing reoccurring constructions, the reduction of the NCBKFLP gain and the increased involvement of cortico-cerebellar processes minimizes reliance on the temporo-parietal Memory (section 2.3.4). A result would be the decay of the constraints imposed to these constructions by default inheritance from their corresponding schemata. The decay of such constraints is accompanied by a range of formal and semantic-pragmatic phenomena. Some of these phenomena are discussed below as reflecting adaptive reductions in cognitive and sensorimotor costs, owing to maximized involvement of cheaper NCB computations.

3.6.3.1. Formal Aspects

The most central aspects of formal changes in grammaticalization involve phonetic attrition and chunking (section 3.2.2). It is interesting to see how both phenomena, either from a perception- or from a production-based perspective, explicitly rely on prediction mechanisms.

3.6.3.1.1. Phonetic Attrition

Phonetic attrition is the macroscopic result of online articulatory reduction. The latter, in turn, has often been treated as the product of automatized language production. This is reflected in Bybee's work, whereby the origins of reduction are found in ''the automatization that comes from repetition of neuromotor sequences'' (Bybee, 2002b, p. 217). From an articulation perspective, studies of CB-induced dysarthria have attracted considerations on the importance of the CB in the acceleration of orofacial gestures,

along with Broca's area (Ackermann & Hertrich, 2000; see also sections 1.3.3 and 5.4.7.2.2 here).

Most importantly, however, phonetic reduction fundamentally relies on the predictability of the reduced form in sentential contexts:

> [r]eduction can be inhibited by the speaker's sensitivity to the predictability of words in the context. If the speaker knows that the word will be easily accessed in the context, because it or related words have already been activated, the reductive automating processes will be allowed to advance. If the word is less predictable in discourse, the speaker is likely to suppress the reductive processes and to give the word a more complete articulation (Bybee, 2007a, p. 243).[58]

Corpus-linguistic and psycholinguistic research concurs that increased ''signal redundancy'' (i.e. predictability) of the articulated item provides the contexts that allow articulatory reduction to occur (e.g. Lieberman, 1963). This is in line with the ''Probabilistic Reduction Hypothesis'', whereby ''word forms are reduced when they have a higher probability'' (Jurafsky *et al.*, 2001, p. 229). Prediction mechanisms in coarticulation effects should come as no surprise. As discussed above, fast, smooth, and compound movements cannot be executed relying solely on external feedback, as there are significant delays involved in both efference and afference. Such movements are made possible by using CB internal feedback (e.g. Wolpert *et al.*, 1998; sections 2.2.1 and 3.5.3 here).

[58] In fact, psycholinguistic evidence suggests that such articulatory reductions occur in speech production without much attention to listener requirements (e.g. Bard *et al.*, 2000), and thus represent a non-attended process.

Similar computations are necessary for perception processes to tolerate such articulatory reductions and to prevent communication from breaking down. Prediction mechanisms thus allow for sufficient comprehension of a degraded-reduced acoustic input, to the extent that phonetic reduction has been considered a ‘‘hearer-based phenomenon’’ (e.g. Haspelmath, 2008, p. 60). The significance of top-down factors in efficient speech perception was discussed in the previous chapter. Phoneme restoration provided a characteristic example of a general, adaptive property of the perception system to employ available information to map poor or missing information onto the intended speech input (sections 2.4.3 and 2.6.3.1). This is why articulatorily reduced lexical items are comprehended with relative ease within constraining sentential contexts, where predictive, top-down processing is invited (Samuel & Troicki, 1998). The capacity of the perceiver to tolerate such degraded speech input was proposed in the previous chapter to rely on the perceptual enhancements that internal models may contribute to in low NCBKFLP gain conditions (section 2.4.3).

3.6.3.1.2. Chunking

Chunking involves the same computations in both language perception-comprehension and production. In grammaticalization, it manifests itself in syntagmatic coalescence, involving cliticization, affixation, or univerbation (section 3.2.2 for examples). The basic principle of chunking is summarized in the ‘‘Linear Fusion Hypothesis’’ whereby ‘‘items that are used together fuse together’’ (Bybee, 2002c, p. 112). In other words, chunking is a process whereby

> […] a frequently repeated stretch of speech becomes automated as a processing unit. The original internal structure becomes less important and can be obscured by phonological change, making the unit more efficient to process […] (Bybee & Scheibman, 1999, p. 577).

At a generic level, chunking is a characteristic sign of automatized processes; it involves individual actions merging with other, spatiotemporally contiguous ones, into a larger, more efficiently performed unit (e.g. Boyland, 1996), as in reach-to-grasp moves, throwing a ball, or writing. Chunking was discussed above as a 'par excellence' CB process (section 3.5.3). In the context of motor output-based metaphors of NCB cognitive functions (see section 2.3.1), NCB structures have been discussed as important for the storage and production of prefabricated, rote-memorized, idiomatic sequences, as opposed to propositionally composed ones. Cases of the former have been thought to include mnemonic sayings, nursery rhymes, songs, childhood jingle, recitation of the alphabet, or even multiplication tables (e.g. Thach, 1997, 2007):

> [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).

The discourse of children with Williams' syndrome (Bellugi *et al.*, 1990), as opposed to that of children with Down's syndrome and CB disease, has provided an elementary dissociation adding weight to the idea above (Thach, 1997, 2007). The former, exhibiting an atrophic cerebrum with a spared CB, excel in their fluent, socially extrovert discourse, which is, however, ''devoid of propositional content'' (Thach, 2007, p. 166). The latter, on the contrary, show deficits in speech development, yet relative sparing of other intellectual activities. Indeed, maybe corroborating the above, NCB activations have been found in an fMRI study of a task involving silent recitation of the months of the year, i.e. a highly practiced speech routine (Ackermann *et al.*, 1998).

However, the linguistic insights gained from such motor-based metaphors of CB contributions are quite limiting (section 2.3.1). Furthermore, there is no reason to assume that these 'non-propositional' contributions are non-semantic or non-linguistic in nature, as suggested in the passage above. The verb-generation impairments in CB pathology are semantic in nature (section 2.4.4.5), and the same type of routinization is arguably involved in inferential processing (see section 3.6.3.2.2 below). Moreover, the organization of motor and cognitive output into larger units may be alternatively explained as a by-product of reliable state estimation of the sensory and cognitive consequences of the action produced (sections 2.2.2 and 3.5.3). In the domain of language, this would mean that such 'idiomatic' modality is equally involved in perception-comprehension and articulation-production.

Indeed, the importance of a predictive component for efficient perception-comprehension is exemplified by the idea of ''automated chunks'' for hearers:

> [t]he boy had learned cannon only in the context of the compound cannonball and that was the only context in which he could access the word. Hearers have automated chunks as well, with analogous priming effects. In the US, upon hearing supreme, one can expect court as the next word; or upon hearing sesame one can expect street (Bybee, 2002c, p. 112).

Such 'hearer-based chunks' and their ''analogous priming effects'' precisely reflect the associative relations between items, established by their frequent temporal contiguity, and provide the perception-comprehension equivalent of CB idiomatic speech output. The generation of the lexical expectations involved was argued in the previous chapter to fundamentally rely on the neodentato-thalamo-cortical component of the NCBKFLP (sections 2.3.4.3 and 2.4.4). The first TMS study reported in this thesis provides some first evidence for selective NCB involvement in such idiomatic phrasal associates, e.g. 'gift-horse', or 'pigeon-hole' (chapter 5).

From another perspective, the involvement of prediction mechanisms in phenomena of syntagmatic coalescence is arguably reflected in the trade-off between ‘‘language redundancy’’, i.e. predictability of items from multi-level linguistic factors and ‘‘acoustic redundancy’’, i.e. predictability based on acoustic salience, in signalling the boundaries between words. According to the ‘‘Smooth Signal Redundancy Hypothesis’’ (Aylett & Turk, 2004), prosodic structure controls the phonetic realization of language redundancy, in a way that recognition likelihood is distributed evenly in utterances. This is achieved by an inverse, complementary relationship between language and acoustic redundancy. In a recent extension of the theory, it has been proposed that the acoustic redundancy of words may be manipulated by signalling their boundaries, and that the occurrence and strength of these boundary markers correlate inversely with language redundancy. Thus, if a word sequence is unpredictable, speakers will be likely to signal the boundary between the words, while frequently co-occurring items, such as subject pronouns and verbs, would not require boundary signalling. In the latter case, the acoustic redundancy provided by the phonetic segments of the words and the relative prominence of their component syllables would suffice for word recognition (Turk, 2010). The way in which such predictions are generated and adaptively modified in the neocerebello-cortical circuits was explained in the previous chapter (sections 2.3.3 and 2.3.4).

The significance of prediction in word-boundary signalling sheds light on the strong correlations noted between ‘‘grammaticalizability’’ of an expression and the occurrence of its grammatical units in a single intonation unit. Constructions that are subject to grammaticalization occur in single intonation units significantly more frequently than their non-grammaticalizable counterparts (Croft, 1995). This is the case for verb-complement structures in English that evolve into auxiliary structures, or into bound tense, aspect and mood inflections in other languages, as the true modal auxiliaries in (19a) or the quasimodals in (19b), or the appearance verbs in (19c) below. This correlation also holds for coordinate structures in English that develop into serial, compound and affixed constructions in other languages, like the motion-action event

coordination in (19d). This adds credibility to the hypothesis that the historical shift from two grammatical units occurring consistently in two intonation units to two grammatical units occurring consistently in a single intonation unit may be another criterion of grammaticalization (*ibid*). In summary, then, language redundancy within a sequence of grammatical units would translate into low NCBKFLP gain conditions in processing the particular sequence. Increased predictability would, in turn, attenuate word-boundary signalling. This would allow the production of that sequence within a single intonational unit, which would in turn pave the way for its grammaticalization.

(19a) would have, should have

(19b) go Ving, start to V, end up Ving, be supposed to V,

(19c) look like NP, seem to VP

(19d) 'he goes and takes the hat to the kid' (Croft, 1995)

If anything, examples like those in (19) demonstrate that spontaneous speech is in fact replete with formulaic and semi-productive phrases, and that formulaicity is far more pervasive in language use than CB neuroscientists envisage in discussing childhood jingles and multiplication tables. This is because prefabricated utterances are accessed without the activation of a high-level schema and provide ''a short cut in processing'' (Wray & Perkins, 2000, p. 15). While constructional schemata allow for the use of novel expressions and the productivity of grammar, the production of novel expressions is computationally costly, involving processing decisions that language users have to make online under time pressure. On the contrary, prefabrications have been computed so often that processing decisions occur with minimal effort. By reducing the amount of utterance planning and sentence processing, attentional resources can be allocated to other communicative aspects (Diessel, 2004). This is why prefabrications are selected, whenever available. For instance, expressions such as 'I don't know __', 'I don't think __', and 'why don't you __' in (20a-c), because of their extremely high frequency, have become prefabricated chunks, and they are produced

and comprehended without a constructional schema (Bybee & Scheibman, 1999). This is manifested in the fundamental differences of these constructions from the less frequent ones they used to share categorical status with. In the former, the pronunciation of 'don't' is reduced, and, while the clauses are formally negated, they do not actually negate a proposition: 'why don't you_' marks a suggestion; 'I don't know_' expresses the speaker's uncertainty or polite disagreement; 'I don't think_' expresses an epistemic stance toward an associated proposition. On the other hand, their less frequent categorical coordinates, as 'we don't eat __' in (20d) are pronounced with an initial stop and a full vowel and serve as negation markers.

(20a) S: no salad? O will you finish the salad darling?
O: I don't know. not now.

(20b) '…I'll drink it but, I don't think about taking it.'

(20c) 'I said why don't you sit down…' (Bybee & Scheibman, 1999)

(20d) We don't eat potatoes.

In this respect, formulaic utterances provide a characteristic case where cortico-cerebellar computations may override and/or bypass cortico-cortical ones in processing an utterance without resorting to the temporo-parietal Memory component. These routinized constructions will thus be computed independently of the structural information they inherit from their corresponding schemata in the temporo-parietal 'construct-i-con' (section 2.3.4). After frequent repetition of the processing of such constructions in dialogical contexts, NCB internal models may accurately acquire the dynamics of processing their phonological-syntactic and semantic-pragmatic properties. The NCB models can then provide the constrained search space for the controller-Unification component to process such formulae on the basis of the co-occurrence of their constituents and according to the interpretational routines associated with them (see next section).

3.6.3.2. Semantic Aspects

In grammaticalization, the shift from 'lexical' to 'grammatical semantics' regularly predates the formal phenomena discussed above (e.g. Givón, 1991; Nicolle, 1998; Kuteva, 2001), or occurs in tandem with them (e.g. Bybee *et al.*, 1994; Bybee, 2007b).[59] Disappointingly, though, despite the claims that grammaticalization constitutes a multi-level reductive change (section 3.5.2), its core semantic component is hardly discussed as a product of these reductions. Instead, semantic bleaching is the aspect promoted as such (e.g. Bybee, 2003), while, at the same time, metaphorical extension and semantic bleaching have been seen as non-central mechanisms for grammaticalization (Bybee *et al.*, 1994; Haspelmath, 1999; section 3.4.3 here). The appeal of semantic bleaching is perhaps better understood in the light of its resemblance to the well-studied processes of cultural ritualization and biological habituation (Haiman, 1994; see section 3.5.2 here). However, it is also because other types of automatization-based semantic change remain elusive. In what follows, I consider some cognitive/ psycholinguistic candidates for this change in the light of NCB computations.

3.6.3.2.1. Cognitive Backgrounding

From a cognitive and discourse-oriented perspective, Boye and Harder (2009) map a widely used in cognitive science ''foreground/ background'' distinction onto that of ''primary/secondary'' in discourse processing. They argue that grammaticalized categories reflect ''coded secondary information status'', and come about as ''the diachronic result of persistent use of a given meaning as secondary information'' (*ibid*, p. 34). The change in the ''structural status'' from ''lexical'' to ''grammatical'' is thus motivated by that in ''usage status'', from ''primary-foregrounded'' to ''secondary-backgrounded'' (Boye, 2008, August; Boye & Harder, 2009). In sentence processing,

[59] Suggestions of the reverse pattern are much rarer (e.g. Lightfoot, 1991); see Nicolle (2007) for discussion.

two pieces of information are seen as competing for foregrounded status, with the loser becoming backgrounded as secondary information. For example, in (21a), the competition is resolved with the information yielded by 'going' winning the competition and preserving primary, foregrounded status. In (21b), the information of 'going' loses in competition, acquiring secondary, backgrounded status. It is precisely in cases like (21b) that the foundations of grammaticalization are found.

(21a) He is going (in order) to fix the room

(21b) He is going to fix the room (tomorrow) (Boye, 2008, August)

Considerations of neuroscientists on the possible NCB contributions to thought processing involve strikingly similar terms and concepts. The NCB has been proposed to adaptively provide a mental ''background'' in contexts requiring extensive processing by a cognitively attended ''foreground''. A characteristic example used in the literature is that of planning out a successful chess strategy: the skilled player would have the problem, the intended strategy, and future moves of the chessmen in the attended foreground, employing the memory of moves below the level of awareness. While this information would be stored in the cerebrum, when planning a winning strategy, it would be fed through the CB as a ''subconscious mental subroutine'' to the foregrounded conscious processes (Thach, 1997, 1998). Moreover, the significance of the CB in supporting dual-task performance, as discussed above, is also quite suggestive. In such paradigms, one of the two tasks gains backgrounded status with the help of the CB, so that attentional resources can be allocated to the foregrounded one (section 3.5.3). Efficient discourse processing would similarly require adaptive NCB backgrounding of certain information, and allocation of cortical attentional resources to the foregrounded one. These CB-based optimizations in discourse processing would initiate the change in usage status of an expression, which in the long run motivates the change in coded structural status.

3.6.3.2.2. Pragmatic Routines

The cognitive-semiotic properties of metaphor and metonymy add another point where grammaticalization changes meet NCB computations. As discussed above, semantic change by metaphorical extension broadly pertains to lexical semantic change, and occurs in the initial stages of grammaticalization (Bybee *et al.*, 1994; section 3.4.3 here). On the contrary, the semanticization of ‘‘pragmatic associations that arise in the flow of speech’’ (Traugott & Dasher, 2005, p. 58), termed ‘‘inference’’, or ‘‘(conceptual) metonymy’’[60] (Hopper & Traugott, 1993), has often been emphasized to play the central role in the evolution of grammatical meanings (Bybee *et al.*, 1994; Nicolle, 1998; Bybee, 2007b).

At this point, I would like to argue that the cognitive properties of metonymy allow the involvement of NCB computations, whereas those of metaphor are selectively suitable for cerebral computations. To begin with, metonymy represents ‘‘semantic transfer through contiguity’’, and is ‘‘indexical’’, while ‘‘metaphor is semantic transfer through a similarity of sense perceptions’’, and is analogical and iconic (Antilla, 1989 [1972], pp. 141-2). As discussed in the previous chapter, the CB does not have the capacity of the cerebrum to store representations on the basis of shared features and similarities. Rather, it stores predictive relations between temporally contiguous events (section 2.3.4.3). This immediately promotes it as a strong candidate for acquiring and processing an indexical sign,[61] or, in other words,

[60]‘Metonymy’ primarily refers to changes based on ‘non-linguistic’ contiguity: e.g. Latin *coxa* ‘hip’ > French *cuisse* ‘thigh’. However, it is extended to contiguities in the conceptual-inferential domain, e.g. *since* (temporal precedence) > *since* (causation). See Hopper and Traugott (1993, pp. 81-2) for discussion of the term and further examples.

[61] Pairings of phenomena and their physical consequences are characteristic examples of indexical signification: smoke-fire, clouds-rain, fever-illness, odor-disintegration.

> a sign, or representation, which refers to its object not so much because of any *similarity* or *analogy* with it, nor because it is associated with *general characters* which that object happens to possess, but because it is in *dynamical* […] *connection* both with the *individual* object, on the one hand, and with the senses or memory of the person for whom it serves as a sign, on the other hand. […] Indices may be distinguished from other signs, or representations, by three characteristic marks: first, that they have *no* significant *resemblance* to their objects; second, that they refer to individuals, *single units*, single collections of units, or single continua; third, that they direct the attention to their objects by *blind compulsion* (Peirce, 1931-58, Vol. 2, pp. 305-306; *italics* mine).

Arguably, then, temporal contiguity and 'non-generalizability' are defining properties for both indexical signalling and CB processes. The fact that ''[i]t makes no difference whether the connection is natural, or artificial, or merely mental'' (Peirce, 1931-58, Vol. 8, p. 368) provides the ground for the extension of these mechanisms to higher cognitive domains. This is in parallel guaranteed by the cytoarchitectural homogeneity of the CB and its massive reciprocal connectivity with all major subdivisions of the central nervous system, which promote a uniform, domain-general CB computation (section 1.3.2).

To begin with, in the transfer from the source to the target concept, inferential processes hardly even involve any shifts in semantic domain; this is because the change does not occur on the grounds of shared categorical properties between source and target:

> […] inferential change […] does not resemble change by metaphor in the least. It is not even clear that it represents a change from a more concrete domain to a more abstract domain […] More importantly, there appears to be no way in which this semantic change can be regarded as a change that transfers an image-schema structure from one domain to another (Bybee *et al.*, 1994, p. 289).

Instead, metonymic transfers involve the acquisition of strong temporal contiguities of interpretational steps. These need to frequently co-occur in order for the inferred meaning to start competing with the conventional one for its subsequent semanticization. For instance, in the transition from Old to Modern French, the unstressed marker of negation 'ne' was supplemented with an emphatic particle 'pas' (originally meaning step, Latin 'passus'). The frequent co-occurrence of the two led to the conventionalization of the Modern French construction 'ne…pas', and to the further grammaticalization of the construction, with the dropping of 'ne' and the metonymic absorption of the negation component by 'pas' (Bybee *et al.*, 1994; Traugott & Dasher, 2005). This non-categorical, associative, frequency-conditioned nature of metonymy offers ideal grounds for NCB involvement. By repeated exposure to the dynamics of the inferential chains in the interpretation of particular constructions, a linguistic NCB model may acquire, via trial and error, the input-output relations of such processes. Upon perception of the particular construct, it may rapidly preactivate the representation of the terminal inferential step; at the same time, it may attenuate the much slower and costlier exploration of temporo-parietal cortex that a step-wise inferential process would involve.

On the contrary, metaphor operates on the basis of similarity and analogy, i.e. by featural overlaps between concepts belonging to different semantic domains. For example, the feature of 'frontness' in the representation of a body part like 'mouth' allows the metaphorical mapping from the 'body' (source) domain to the 'space' (target)

domain (see example 3, section 3.2.2). The CB does not store categorically organized representations, and is thus unable to compute the similarities between subordinate terms of different categories on the basis of such featural overlaps. Mechanisms of similarity-based learning were shown in the previous chapter to reflect cortical algorithms of unsupervised learning (section 2.3.4.3). Furthermore, metaphorical extension has been seen as a process that, unlike metonymy, may occur abruptly (Bybee *et al.*, 1994). On the contrary, it is only through repeated exposure to the dynamics of a certain process that the CB is trained (section 2.2.2). Thus, by its very nature, metaphor is less suited for NCB involvement. Furthermore, while metaphor involves the quite innovative creation of a novel form-meaning mapping (and thus a 'competence change'), 'metonymy' arguably reflects a shortcut behavior in the inferential process: the stored input-output sequences may override the very process by which these outputs are yielded. This distinction would include metonymy in the 'performance changes' proposed to comprise grammaticalization (sections 3.4.3 and 3.4.4 above).

This shortcut behavior brings the discussion to the notion of ''pragmatic routines''; these represent the ''[…] tendency of the mind, in its search for relevance, to use assumptions and inferential processes which have been used in processing the stimuli on previous occasions'' (Vega Moreno, 2007, p. 117). Re-occurring pragmatic inferences may become so accessible that minimal effort would be required for their execution. Thus, pragmatic routines would operate by ''compiling or short-circuiting the inferential steps involved in comprehension'' (Vega Moreno, 2007, p. 117), directly reflecting NCB computations. In other words,

> [t]he more familiar the hearer is with a particular combination of words […] the more likely is that the most highly activated assumptions will be those which had been considered or derived in processing the string on previous occasions (*ibid*).

‘Familiarity’, translated into low NCBKFLP gain in processing, would involve maximized NCB involvement in determining the output of the emulated process. In the case of a precompiled construct, interpretation would be biased to output the same assumptions as those generated in previous occasions of its processing, with little access to the representations of other concepts involved in the inferential process. Such cases would fall under the semantic-associative computations held to rely on the NCB in the previous chapter (sections 2.3.4.3 and 2.4.4.5). The third TMS study reported here (chapter 7) also provides some first evidence for NCB involvement in semantic-associative relations in the comprehension modality.

3.6.4. Micro-Reanalysis and Cortical Schematization

The discussion now briefly turns to the notion of ‘cortical schematization’ as a necessary component for ‘CB-optimized’ constructions to gain categorical status and ultimately give rise to new schematic constructions.

3.6.4.1. Cortical Schematization in Acquisition

In functionalist grammaticalization theory, child language acquisition has been proposed to contribute to the coding and the productivity of a grammaticalized variant via processes of ‘‘micro-reanalysis’’ (Haspelmath, 1998). By such process, the diachronically elder variant is marginalized; the newer, grammaticalized one, on the basis of its higher frequency in the learning set the infant is confronted with, acquires basic, ‘unmarked’ status, and becomes productive. Hooper (1976) independently employs the same mechanism to describe the conventionalization of phonetically reduced variants (section 3.4.4 here):

> [a]fter the initial stages, the language acquisition process begins to play a role, for it is probably in the transmission of the language that restructuring takes place. A very frequent word such as *every* may be variable for an adult speaker, but a new learner may take the schwa-less pronunciation to be the norm or base form. The younger speaker will have a schwa-less underlying form, and the process will be complete for that word. The change in underlying forms will progress similarly over several generations (Hooper, 1976; reprinted in Bybee 2007, p. 31).

For the grammaticalizing construction to acquire schematic status, analogical mechanisms of rule generalization should be operand. For instance, in the example of the procliticization of object pronouns in Spanish above (18), it would be only particular sequences of verbs and co-occurring pre-verbal object pronouns that become routinized into the same intonational contour. Apparently, this construction could not have developed as a product of routinization for each separate 'object pronoun-verb' sequence- some of these combinations are simply far from frequently used in dialogue. Instead, the highly entrenched ones, argued here to involve NCB circuitry in their processing, would form the foundations for the formation of a new categorical representation of a schematic construction, on the basis of which a much larger set of verbs and pronouns may be processed. The same generalization mechanism would be seen in the development of the Romance perfect. In (22a), agreement between the object ('vos') and the adjectival participial ('fatigatos') is overt. In the perfect tense construction, though, in (22b), agreement between the object ('omnia') and the participle ('probatum', instead of 'probata') has decayed. The construction formed in (22b) was apparently derived by instances of the construction in (22a) where the object, and, via agreement, the participial as well, were in neuter singular form. After this decay in agreement, the generalization that occurred in the object slot from neuter singular

contexts to those of all numbers and genders in the perfect tense construction (22b) was a product of ‘‘analogy’’ or ‘‘rule generalization’’ (Hopper & Traugott, 1993).

(22a) Metuo enim ne ibi vos habebam fatigatos
Fear.1.SG for lest there you: ACC.PL have-1SG tired-ACC:PL
‘For I fear that I have tired you.’

(22b) Haec omnia probatum habemus
Those:ACC:PL all-ACC-PL tried-PART(?) have-1PL
‘We have tried all those things.’

(Fleischman, 1982, p. 120; found in Hopper & Traugott, 1993, pp. 57-8)

For this to occur, I would like to suggest that mechanisms of schematization in child language acquisition are necessary. The latter involves an analogical learning process by which categories are formed on the basis of their featural overlaps (Langacker, 2000; Diessel, 2004). As argued in the previous chapter, schematization is a function of cortically-based, unsupervised learning algorithms, and cannot be accommodated by CB supervised ones. This is corroborated by findings involving Broca’s area (and not the CB) in the extraction of schematic representations of letter strings in artificial grammar learning paradigms (sections 2.3.4.3 and 2.4.4.4). A characteristic instance of schematization is the over-regularizing behavior of infants, which is a powerful driving force for reanalysis and analogical levelling. As discussed above, those two phenomena are described as ‘‘categorical’’, and are contrasted with the ‘‘gradual’’ nature of grammaticalization (Lehmann, 2004; section 3.4 here).

The distinction between ‘schematization’ and ‘routinization’ promoted here, and that between ‘cortico-cortical’ and ‘cortico-cerebellar’ processing, respectively, are compatible with the ‘‘analytic’’ vs ‘‘formulaic’’ distinction promoted in Wray and Perkins (2000). In the initial ontogenetic phase, where the analytic processing modality prevails, the cognitive apparatus of language-acquiring infants

> […] identifies, selects and stores a sufficient and requisite number of salient formulaic linguistic items to activate a specifically language-oriented analytical mechanism which, through identifying commonalities among the stored formulas, begins the process of creating a generative grammar for the language of the child's environment somewhere between 20 and 30 months of age. The process continues until roughly the age of 8, and this stage of development is marked by a preference for analytic over formulaic language processing (*ibid*, p. 21).

After the 8th year of age, the holistic modality becomes heavily involved in carrying out the processing bulk in casual linguistic interaction. In doing so, it offers a range of ‘‘processing short-cuts’’, which are ‘‘crucial to managing an over-demanding on-line processing system’’ (*ibid*, p. 17). In such cases, comprehension

> […] becomes a top-down pragmatically driven process of formulaic `macro-processing', with the bottom-up grammatical `microprocessing' mechanism only being used as a default in cases where macroprocessing fails to yield a sufficiently relevant interpretation (*ibid*, p. 21).

Interestingly, the CB shows a delayed postnatal developmental maturation compared with the cerebrum (e.g. Hossain *et al.*, 2004), and it does not reach its full growth until the 15th to 20th year of age (Leiner *et al.*, 1986). While research on the maturation of particular NCB lobules is underresearched (C. Miall, personal communication, 24 June 2010), this general pattern adds weight to the idea of cerebral analytical computations of schematization preceding the CB holistic ones of routinization.

3.6.4.2. Cortical Schematization within Single Cohorts

But would the micro-reanalysis of newly introduced grammaticalization variants rely on schematization exclusively in the language acquisition phase? An intriguing possibility is offered by the notion of 'forced CB-to-cerebrum learning', whereby, in the long run, the CB trains its cortical output loci, making the representations in the latter more efficient:

> [o]n the one hand, these [CB-induced] alterations would serve to modify the collective computations being performed by the cortical network […] On the other hand, the same alterations would promote changes in the weights of the Hebbian synapses on pyramidal neurons, that, in the long run, would move the network's attractors closer to the points being forced by cerebellar and basal ganglia modifications. As a consequence, the frontal cortex would become trained to perform, in a highly efficient and automatic fashion, those particular functions being forced on it by its subcortical [CB and BG] inputs (Houk & Wise, 1995, p. 106).

Unfortunately, direct evidence is not yet available, though the plausibility of this idea is implied in assumptions of cerebello-cerebral interactions. For example, the changes in cerebral metabolism consequent on CB cortical lesions, e.g. in crossed cerebro-cerebellar diaschisis (section 1.3.3), imply that reduced CB outflow affects long-term cerebral processing. But if the cerebrum receives modified CB inputs in the longer term, it would be highly surprising if there was no change in the cerebrum to reflect that (C. Miall, personal communication, 22 December 2009). Suggestively, in a recent study, individuals with damage to their left, dominant for speech, cerebral hemisphere, activated, during recovery, their previously non-dominant, right cerebral hemisphere, along with the crossed left CB hemisphere. The new patterns of functional connectivity

that emerged were proposed to result from the CB assisting the right cerebral hemisphere to subserve mental activity (Connor *et al.*, 2006; see Thach, 2007, for some discussion).

Conceivably, then, the associative representations of particular constructions developed in the CB might be copied back to the cerebral language-related loci. The hypothesis would be that the cerebrum may then systematize and schematize these CB-acquired associative copies. In that way, traces of new categories would be made possible within a single cohort of adult speakers. For instance, in (23) below, the cerebral memory representations (C) of the construction ([X[YZ]]) in (23a) below are instantiated, among others, with the frequently occurring 'constructs' (physical instantiations of particular constructions) $[x_2y_1z_3]$ and $[x_2y_1z_4]$; 'X, Y, Z' would represent categories of items, and 'x, y, z', would represent members of such categories. CB memory (CB) may acquire the dynamics of formal and semantic processing of these constructs, and store associative representations that bear no categorical, and thus no structural information of the items involved (23b). By being processed on the grounds of their formal co-occurrence and their routine interpretation, these constructs develop properties absent from their schematic category and from their categorical coordinates. Via forced CB-to-cerebral learning, such associative representations may be acquired in turn by the cortical memory component (23c). This may provide input for the construction of new categorical representations, by the analogy-driven schematization processes, thus giving rise to new schematic constructions, on the basis of these constructs $[x_2y_1Z']$, with new formal and semantic properties. By further analogical levelling, the paradigmatic breadth of Z' may expand to the much larger Z (23e), allowing a greater range of lexical items to fill the slots of the new construction.

(23a) C = {[X[YZ]]} CB = { }

$X = \{x_1, x_2, x_3, ..., x_p\}$
$Y = \{y_1, y_2, y_3, ..., y_n\}$
$Z = \{z_1, z_2, z_3, \ldots, z_m\}$

(23b) C = {[X[YZ]]} CB = $\{[x_2y_1z_3], [x_2y_1z_4]\}$

(23c) C = {[X[YZ]], $[x_2y_1z_3]$, $[x_2y_1z_4]$} CB = $\{[x_2y_1z_3], [x_2y_1z_4]\}$

(23d) C = {[X[YZ]], $[x_2y_1Z']$} CB = $\{[x_2y_1z_3], [x_2y_1z_4]\}$
$Z' = \{z_3, z_4\}$

(23e) C = {[X[YZ]], $[x_2y_1Z]$} CB = $\{[x_2y_1z_3], [x_2y_1z_4]\}$

3.7. Basal Ganglionic Computations in Grammaticalization

As already suggested, particular brain structures are only relevant with grammaticalization to the extent that they implement computations that support the psycholinguistic mechanisms underlying grammaticalization (3.3.3). Hence, while NCB circuitry was argued above to hold a central place in such grounding, other aspects of grammaticalization may involve different psycholinguistic processes; these may employ different computations that may be performed by different brain structures. In Argyropoulos (2008a), I had suggested that both CB and basal ganglionic computations are involved in grammaticalization changes. In fact, in a number of neuropsychological models, the basal ganglia are shown to contribute, along with the CB, to multimodal automatization processes. In the context of motor automaticity, for instance, the CB has been proposed to associate experiential contexts with compound motor responses, whereas the basal ganglia to fluently gate competing alternative responses by selecting the most appropriate ones (Thach *et al.*, 1993). The strong relationship between automatization and grammaticalization (section 3.5.2) thus encourages the investigation of the role of this structure in language processing. Briefly, here, I present how the basal ganglia may be involved.

Grammaticalization studies have offered insight into the cognitive representation of constructions (Bybee 2005, p. 10), and, importantly here, into the categorical gradedness of grammatical representation. Usage changes are concomitant with changes in cognitive representation, and thus more flexible and gradient grammatical categories are required (e.g. Givón, 1979; Haspelmath, 1998; Bybee 2007b, pp. 973-4). For instance, the ‘‘verbality/ prepositionality’’ scale in English ranges from less grammaticalized, verb-like items, e.g. ‘preceding, facing, considering’, to more grammaticalized, preposition-like ones, such as ‘during, pending, past, ago’ (Kortmann & König, 1992, p. 684; see discussion in Haspelmath, 1998).The desideratum of

gradedness has recently attracted researchers from computational and psycholinguistic probabilistic modelling (Zuraw, 2003 for a review), encouraging the definition of categoricity on the basis of the particular constructions that each item occurs in (Croft, 2001; Goldberg, 2004). Similarly, in Pulvermüller's (2002) neuronal syntax, lexical categories are defined by the set of the very complements lexical categories require, i.e. by their ''sequence regularities''.

If assignment of categorical status to a particular item relies on recognition of the particular constructional context, then, in online processing, the representations of these contexts may be activated once the item is encountered. These representations would assign different candidate analyses to the particular item, and their competition may be resolved with subsequent input. In (24a-b), for instance, the verb representation of 'considering' competes with its prepositional representation, and disambiguation may require further input, even beyond the sentential context. A necessary component of an efficient parser thus becomes the gating of competing structural analyses. One such mechanism may be found in recent lexicalist processing models, where competing analyses send each other inhibitory signals, reducing the competitor's strength (Vosse & Kempen, 2000; Hagoort, 2003).

(24a) Considering retirement, I still have years to come.
(24b) Considering retirement, I called my manager.

Such inhibitory processes among competing variants may be undertaken by cortico-striatal circuits. This is indeed reflected in what Pulvermüller (2002) en passant calls ''striatal regulation of cortical activity'', operating in cases where more than one of these ''sequence regularities'' become available for an item online. Recent evidence implicates basal ganglionic inhibition in sentence processing. Subjects with Tourette's syndrome have shown speeded processing of procedural (both linguistic and non-linguistic) knowledge, which has been attributed to their basal ganglionic abnormalities in the inhibition of frontal cortical activity (Walenski *et al.*, 2007). Correlations have been found between sentence comprehension and Stroop task performance in

Parkinsonians (Grossman *et al.*, 2002), while their compromised capacity of parsing relative clauses has been attributed to ‘‘deficits in cognitive set-switching’’ or to ‘‘underlying inhibitory processes’’ (Hochstadt *et al.*, 2006). In general, inhibition and reinforcement are taken to underlie probabilistic representation: basal ganglionic patients exhibit deficient probabilistic category learning (Knowlton *et al.*, 1996), while in the acquisition phase of such tasks, striatal activation is involved for normal individuals (Poldrack *et al.*, 1999). Basal ganglionic activation in the comprehension of syntactically ambiguous sentences has also been discussed as reflecting a process of ‘‘unchoosing’’ the initial syntactic commitment upon reception of incompatible syntactic input (Stowe *et al.*, 2004).

At the intra-generational level, then, changes in the distributional patterns of particular items would be efficiently tracked by a constant, dopamine-mediated regulation of the inhibitory strengths among competing syntactic contexts. This fuzzy view of grammatical categoricity would translate into sets of item-particular ‘cortico-striatal signatures’, constantly updated throughout the lifespan of the individual (figure 3.3).

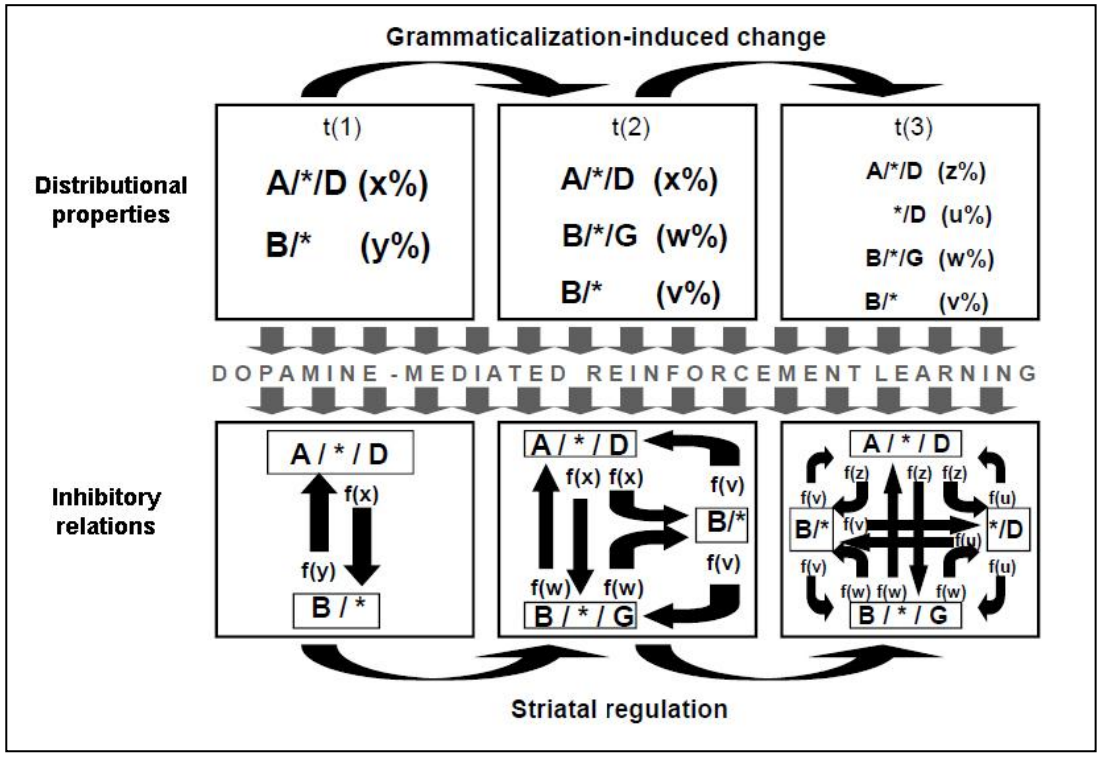


Figure 3.3: Cortico-striatal negotiation of linguistic categoricity. Cortico-striatal circuits regulate the competition between candidates for unification in online sentence processing. Changes in the distributional properties of linguistic items are efficiently translated into changes of inhibitory strengths among the competing alternatives that the item unifies with in processing. A, B, C, D, G: preceding and following morphosyntactic environments of linguistic item (*); x. y, z, w, v, u: values of the frequency in which a particular item (*) occurs in a particular environment. The function 'f' translates those distributional strengths into strengths of inhibitory signals that the representations of the different morphosyntactic constructions send to one another in online processing; t(1,2,3): different arbitrary stages in the ontogenetic (intra-generational) timescale. Figure adapted from Argyropoulos (2008a, p. 15) with permission © 2008 World Scientific.

3.8. The Cerebellum, Broca's Area, and Grammaticalization

In his pioneering work bringing historical linguistics and neuroscience together, Givón has drawn attention to the significance of Broca's area in the diachronic emergence of grammatical structure.[62] The cortex around Broca's area is involved in a wide range of non-language-particular cognitive functions, such as motor skills, action-planning routines, and complex hierarchic object-combination, visual-tracking, and tool-using routines (e.g. Givón, 1995, pp. 420-1). Comparative surveys of non-human primates and human children support the idea that 'Broca's area as a grammatical processor' reflects the late ontogenetic and phylogenetic extension of pre-linguistic uses of the primary motor cortex for sequential-hierarchic skills (Greenfield, 1991). The idea is thus promoted that grammaticalization as automatization primarily involves increased involvement of Broca's area in communication. However, there are certain issues that need addressing at this point.

Throughout his research, Givón has used the term 'grammaticalization' in a much broader fashion than the way it is employed here and the way typically used in the literature. In particular, Givón's definition is based on the striking similarities among the contrasts between creole and pidgin, adult and child language, or healthy and aphasic adult speech (e.g. Givón, 1979, p.223). Those two modes of communication, described as ''grammaticalized vs pre-grammatical speech'' (Givón, 1989, p.262), fundamentally differ with respect to the information load/speed trade-offs, and reflect differences between automatic and attended processing. The same distinction is preserved in Givón's most recent work, where ''pre-grammatical pidgin before grammaticalized language'' is described as one of the major trends in language ontogeny (Givón, 2009, p. 123).

[62] Needless to say that Givón's ecology has never exclusively committed to a particular brain locus, and is openly pluralistic (see Givón, 2009, for an up-to-date account).

Grammaticalization, for Givón, then, pertains in general to the transition from a pre-/non-grammatical to a grammatical mode of communication and information processing. On the contrary, the definition of grammaticalization adopted here is much stricter: it refers to the evolution of constructions within an already established (cognitively and socially) grammatical mode of communication. The two definitions assumed can actually be seen as radically different from each other: 'grammaticalization' as the introduction of the grammatical mode of communication, in contrast to 'grammaticalization' as the 'overriding/bypassing' of such a mode, with constructions deviating from their categorical properties in online processing because of cheaper emulative computations. Givón's broad definition is exemplified in the emphasis lent in his work to instances of ''syntacticization'', i.e. the development of syntactic complexity out of loose, paratactic speech. The shift of the loose paratactic topic-comment construction (25a) into one of subject-verb agreement (25b) would be such an example. In the stricter construal of grammaticalization, syntacticization represents a ''non-prototypical'' case (e.g. Boye, 2008, August).

(25a) My ol' man, he rides with the Angels.
Topic Comment

(25b) My ol' man he-rides with the Angels.
Subject Agreement

(Givón, 1979, p. 209; cited in Boye, 2008, August)

Above all, this broad definition of grammaticalization as the transition from a pre- or non- grammatical to a grammatical mode of communication would accommodate computations of child language acquisition within those underlying 'grammaticalization'. The present account, on the contrary, assumes a much stricter distinction on the learning mechanisms between child language acquisition and adult language use, in line with work in grammaticalization and child language acquisition studies (e.g. Hooper, 1976; Haspelmath, 1998; Diessel, 2004).[63] In particular, child language acquisition was proposed above to involve the schematization of linguistic

[63] See chapter 7 in Givón (2009) for an account against this distinction.

input, and thus to heavily rely on cortical unsupervised learning algorithms (section 3.6.4). On the contrary, the routinization of language processing in adult dialogical contexts was argued to primarily involve NCB feedforward control, overriding and/or bypassing the slower cerebral categorical processes (sections 3.6.2 and 3.6.3). Future imaging and clinical research should help dissociate between the conditions of the involvement of Broca's area and those of the CB in language processing. Encouragingly, currently available evidence suggests that Broca's area, and not CB circuitry, is significant for artificial grammar learning (sections 2.3.4.3 and 2.4.4.4). On the contrary, the deficits of CB patients in practice-related performance optimization (sections 2.4.2 and 3.5.3) do not reflect any known cerebral impairments. In other words, while both Broca's area and the CB belong to the motor, non-attended processing loop (section 3.5.1), their contributions to these broadly defined 'automatization' operations may arguably be different. Along with further neuropsychological research, more refined cognitive descriptions would also be needed to elucidate the different roles of these two structures.

3.9. Conclusion

In this chapter, the idea of automatization as the cognitive core of grammaticalization provided the necessary bridge between language diachrony and neurolinguistics. The significance of the CB for automatization allowed us to explore the particular psycholinguistic aspects of grammaticalization in the light of the NCBKFLP. In particular, grammaticalization was argued to rely on the adaptive minimization of the NCBKFLP gain, as induced by the routinization of processing particular constructions in casual adult dialogical interactions. Grammaticalization changes were thus discussed as the product of maximized involvement of associative neocerebello-cortical computations in language processing that a low NCBKFLP gain modality invites.

Conclusion of Volume I

In the first, theoretical part of the thesis (chapters 1-3), neocerebellar computations of feedforward control and state estimation were brought together in a synthesis with psycholinguistic work on language perception and comprehension, along with historical linguistic work on grammaticalization processes. Neocerebellar compartments were argued to provide associative (predictive) linguistic computations capable of biasing, overriding, and/or bypassing the cerebral categorical ones. Neocerebellar processes were proposed to underlie a wide range of psycholinguistic and historical linguistic phenomena, thus making possible the empirical investigation of the hypothesis. Volume II presents the first steps taken here towards the latter.

The Neocerebellar Kalman Filter Linguistic Processor: From Grammaticalization to Transcranial Magnetic Stimulation

Volume II

Giorgos P. Argyropoulos
B.A. (Hons.)

A thesis submitted in fulfillment of requirements for the degree of Doctor of Philosophy

to

Linguistics and English Language
School of Philosophy, Psychology and Language Sciences
University of Edinburgh
August, 2010

Volume II

Empirical Chapters

Introduction to Volume II

In the first, theoretical part of the thesis, the neurolinguistic model of the Neocerebellar Kalman Filter Linguistic Processor was developed (sections 2.3.3 and 2.3.4). Within this context, a number of experimental hypotheses were formulated (section 2.6). The thesis now turns to the first steps taken here towards the validation of these hypotheses, employing transcranial magnetic stimulation (TMS) of neocerebellar loci. After discussing the motivation for the particular methodology and the stimulation protocols applied (chapter 4), the thesis goes on to report the three TMS studies conducted, which are the first to combine cerebellar stimulation with linguistic tasks (chapters 5-7). Despite the considerable technical limitations, these three studies provide considerable evidence in support of the theoretical work presented in the first part (chapter 8).

Chapter 4

Experimental Methodology

> ‘‘As any schoolboy with a toolkit or a broken toy soon appreciates, to find out how a machine works you need to take it apart, and to put it back together again, you need to know how it works. The next lesson is that, no matter how hard you try, you always end up with a handful of leftover nuts and bolts. These remaining components can be informative: will your machine still work without them?’’ (O’ Shea & Walsh, 2007, p. 196).

4.1. Introduction

This chapter provides a brief introduction to the general features of the transcranial magnetic stimulation (TMS) framework, and, subsequently, the reasons why TMS experimentation was preferred over clinical studies for the empirical investigation of the hypotheses formulated in the first part of the thesis. The chapter then continues with the explanation of the methodological parameters pertaining to the studies conducted here, which involve the first linguistic CB TMS experiments in the literature.

4.2. Introduction to TMS

A comprehensive introduction to TMS would be beyond the scope of the present thesis; the reader may instead refer to the introductory articles to TMS cited in this section.

TMS is a non-invasive technique for the electrical stimulation of neural tissue in conscious human subjects. According to Faraday's principle of electromagnetic induction, when a pulse of electrical current is introduced through a coil of wire, a magnetic field is generated (Faraday, 1965). If its magnitude changes in time, the magnetic field induces a secondary electrical current flow in a second wire, i.e. a nearby conductor. The rate of change of the field determines the size of the current induced. In TMS, the 'first wire' is the copper-wire, encased in plastic stimulating coil and the 'second wire' is a targeted region of the brain (Pascual-Leone *et al.*, 1999). The stimulating coil is held over a subject's head, and applied over the scalp at a point overlying a specific brain area. When a capacitance is discharged through it by a magnetic stimulator, a brief pulse of large, rapidly changing current flows through its windings, generating a transient magnetic field oriented orthogonally to the plane of the coil. Unlike transcranial electrical stimulation, the magnetic field passes through the subject's skin, scalp and skull with negligible attenuation (Rothwell, 1997), at a depth of approximately 1.5-2 cm beneath the scalp (Rudiak & Marg, 1994; Wassermann, 1998; Keenan & Pascual-Leone, 1999), only decaying by the square of the distance. This time-varying magnetic field induces a much smaller, oppositely directed secondary ionic current perpendicularly to the magnetic field in the subject's brain, flowing tangentially with respect to the skull and inducing a physiological response in the underlying neural tissue. Thus, ''strictly speaking, transcranial magnetic stimulation is a misnomer, as the magnetic field appears to simply represent a bridge between the current in the stimulating coil (primary current) and the current induced in the subject's brain (secondary current). Therefore, TMS might be best seen as a form of 'electrodeless, non-

invasive electric stimulation'.'' (Pascual-Leone *et al.*, 1999, p. 1230). In general, the secondary current induced by TMS triggers activity changes, excitatory or inhibitory, that are effectively random with respect to the organized signals required to perform a task. This transient 'neural noise' is thus introduced to the neural computation being performed; its disruptive effects can be manifested in longer reaction times (RTs), or even errors (Pascual-Leone *et al.*, 1999; Walsh & Rushworth, 1999).

TMS provides a valuable tool for cognitive neuroscience in a number of different ways. Most importantly, ''among the many methods now available for imaging the activity of the human brain, magnetic stimulation is the only technique that allows us to interfere actively with brain function.'' (Pascual-Leone *et al.*, 2000, p. 232). This is because TMS can be employed as a ''lesion technique'' by transiently disrupting focal brain activity and causing ''virtual lesions'' (Walsh & Pascual-Leone, 2003). It can also be used to study the timing of the contribution of an accessible brain region in a behavior. Moreover, it can be combined with functional magnetic resonance imaging (fMRI), single unit recording, and high-resolution electroencephalography, for the investigation of task-related networks and changes in effective brain connectivity. Furthermore, since the neural noise introduced by TMS often leads to transient performance enhancements, it also has promising clinical utility, and, since it can influence brain function if delivered repetitively, it may be employed diagnostically, prognostically, and even therapeutically (Kobayashi & Pascual-Leone, 2003). Unsurprisingly, then, since its introduction, TMS has been extensively used for the stimulation of peripheral nerves and brain tissue involved in a wide range of functions, such as motor control, swallowing, vision, attention, memory, speech and language, as well as in movement disorders, epilepsy, depression, stroke, motor conduction in human development, pain and plasticity.

4.3. Methodological Significance of Cerebellar TMS

I now turn to the methodological advantages of TMS as a lesion technique compared with clinical and other imaging techniques with respect to the study of NCB cognitive functions. Those advantages provide the reasons for which TMS methodology was preferred in taking the first steps for the assessment of the NCBKFLP hypothesis (chapter 2; Argyropoulos, 2009).

Compared with the rest of the human brain-imaging techniques, TMS has the significant advantage of the capacity to demonstrate ‘‘causal structure- function relationships’’ (Sack, 2006, p. 593). A non-invasive mapping technique such as fMRI provides evidence for transient local changes in neural activity during task performance, without, however, proving that the loci activated are actually used for the task. In other words, fMRI would merely provide evidence for an association, but not a causal link between the activated region and the behavior. The combination of TMS and functional neuroimaging provides a novel solution to this problem: activity in brain areas associated with a behavior in fMRI or PET can be disrupted by TMS in order to investigate subsequent alterations in this behavior, and thus assess the causality of these relationships (Paus, 1999). In this fashion,

> TMS provides a unique methodology for determining the true functional significance of the results of neuroimaging studies and the causal relationship between focal brain activity and behavior (Pascual-Leone *et al.*, 1999, p. 1229).

The significant role that TMS could play in exploring CB language functions becomes apparent when seen in the light of the problems in CB neuroimaging studies; a fortiori since the latter have provided the leading paradigm in advocating a CB

involvement in higher cognitive aspects (Desmond & Fiez, 1998; Timmann & Daum, 2007). Imaging results have often been met with skepticism, since changes in CB activation may easily be attributed to involvement in sensorimotor aspects of the task, instead of any putative cognitive ones (Timmann & Daum, 2007). For example, inconclusive evidence persists on the role of CB regions in word reading and word-generation tasks. Comparing activations between verb generation and verb reading, Frings *et al.* (2006) found activations in the right posterolateral cerebellar (PLCB) lobule HVI and Crus I as a measure of verb generation; these were lateral from the paravermal activation of lobule VI, which was associated with speech articulation. Their findings were in line with those of Petersen *et al.* (1989), who had compared noun reading with verb generation. However, in contrast to Petersen *et al.* (1989), no CB activations were found comparing verb generation with noun reading by Frings *et al.* (2006). Furthermore, these PLCB activations were also found as a measure of noun reading in inner speech, and could thus provide counterevidence for CB involvement in higher linguistic functions (Frings *et al.*, 2006). Similarly, in another fMRI study, activation within the same region was found during silent recitation of the names of the months of the year; this was interpreted on the basis of PLCB involvement in timing aspects of both inner and overt speech production, rather than in cognitive operations (Ackermann *et al.*, 2004). In cases like the above, TMS methodology would help dissociate causal relationships from simple associations between CB activations and task performance. Different, neighboring areas could also be targeted, allowing precise, dissociative mapping of a brain area to a particular behavior. This has already been applied, for instance, in studying CB motor functions, by contrastive vermal and hemispheric CB stimulation (Théoret *et al.*, 2001).

Importantly, TMS bypasses some of the most common obstacles posed in CB clinical research. CB patients differ with respect to the depth of the lesion, which might be constrained only to the CB cortex or might extend to the nuclei. The nature and extent of the diseases is often such that it is not possible to draw any conclusions about possible correspondences between lesion location and specific deficit (e.g. Pickett, 1998). Most

studies also involve chronic CB patients, where compensation for and plastic reorganization of the lesioned circuits might obscure the interpretation of findings (Timmann & Daum, 2007). CB-induced language deficits, for example, are quite mild and transient, and might involve significant compensation from cortical language-related areas (Fabbro *et al.*, 2004). On the other hand, while impairment may be more prominent in patients with acute CB damage, patients with selective CB lesions are extremely rare. Even in such cases, the heterogeneity of deficits and lack of specificity of their effects, such as hydrocephalus, depression or global effects on brain metabolism might significantly affect cognitive function (Timmann & Daum, 2007). In fact, most CB groups show a certain degree of extra-cerebellar dysfunction, e.g. brainstem damage and increased intracranial pressure accompanying CB stroke and tumors. Even purer degenerative CB disorders, such as spinocerebellar ataxia type 6, or sporadic adult onset ataxia show mild extra-cerebellar symptoms (Frank *et al.*, 2007). This picture has made neuroscientists to consider data from patients with combined CB and extra-cerebellar damage to be of limited value for the issue of CB involvement in cognition (e.g. Daum & Ackermann, 1997). It has also made researchers turn to the very rare patients with pure CB disorders to determine the specific CB contribution to cognitive impairment, as well as to validate the findings in patients with a less focused disease (Timmann & Daum, 2007).

The restricted subject pool of selective CB lesions, along with the great heterogeneity of the larger non-restrictive ones make the replicability of findings a major issue. For instance, studies reporting deficits in frontal lobe function, visuospatial processing or non-motor skill learning have been found difficult to replicate (Daum & Ackermann, 1997). The debatable presence of a verb-generation deficit in PLCB patients provides a characteristic example. Impairments in verb generation (Fiez *et al.*, 1992; Gebhart *et al.*, 2002) have not been shown in every study with PLCB patients. For example, Helmuth *et al.* (1997), and Richter *et al.* (2004) found effects of dysarthria only. Richter *et al.* (2004) emphasize the small subject pool size in Gebhart *et al.* (2002), as well as the high average age of subjects as confounding factors. On the other hand,

the patient populations in Helmuth *et al.* (1997) and Richter *et al.* (2004) were arguably more heterogeneous, including patients with lesions in broader CB areas, as well as patients with CB degeneration. There was also a difference in the amount of time between the CB damage and the time of testing across the articles: in Gebhart *et al.* (2002), patients with recent CB damage were tested, while Richter *et al.* (2004) involved patients with CB atrophy (A. Gebhart, personal communication, 15 May 2008).

These issues, bound to the properties of 'naturally occurring' CB lesions, do not arise in the case of TMS-induced virtual lesions. TMS studies allow a considerably non-restrictive subject pool, where the experiment may be repeated in the same subject, providing an opportunity for careful, strictly controlled experimental design. TMS studies are also conducted acutely, avoiding ''the specter of neural compensations'' (Lomber, 1999; Walsh & Pascual-Leone, 2003): time is insufficient for functional reorganization to occur during single TMS events, and thus no substantial confounds are expected due to recovery processes (Walsh & Cowey, 2000). Furthermore, TMS is conducted on normal subjects and is more focal than naturally occurring lesions, thus eliminating the confounding factors of the heterogeneity and extension of such lesions seen in patient studies. The fact that focal TMS coils usually affect the first cortical areas of the neural tissue, and, in the case of the CB, the Purkinje cell cortical layer allows more consistent explanations of the effects according to the microcircuitry of the neural area (see section 4.4.9 below).

The subtlety of CB cognitive deficits also makes the selection of well-matched control groups over several variables (e.g. age, IQ, education level) both a methodological necessity and, wherever this is not the case, a confounding factor (Daum & Ackermann, 1997). Potential differences in the premorbid ability of individual patients may further compromise the validity of the results: the patient's performance is compared with the average performance of a control group, and not with their own premorbidly. On the contrary, the reversibility of TMS-induced virtual lesions provides the possibility of employing individual subjects as their own controls: non-TMS trials

are compared with TMS trials in online protocols, or, in the case of offline, distal TMS (section 4.4.3), trials before TMS are contrasted with trials after TMS, in a within-subjects design, avoiding potential inter-subject confounds (see section 4.4.3 below).[64]

Basic processes involved in response production may also be affected in patient populations, either due to dysfunction of the area of interest per se, or due to the extent of the brain damage. This is a recognized problem in CB clinical studies, where the well-established significance of CB structures for eye and finger movements and overt speech, motor impairments means that CB impairment may compromise performance on cognitive tasks, in a rather confounding fashion. Suggestively, it has been argued that the attentional deficits observed in CB patient studies pertain only to the test procedures with increased oculomotor, motor, and/or working memory demands (Timmann & Daum, 2007). Indeed, unlike general intellectual capacities or memory, deficits in motor learning or temporal processing are consistently observed in CB patients (Daum & Ackermann, 1997). For instance, acquired dyslexias and dysarthrias often accompany vermal-paravermal damage, along with dysmetria, rendering response production in the form of button presses or oral responses in the simplest of tasks (providing 'yes' or 'no' responses to a lexical decision task) very difficult; it also often makes the use of a more sensitive measure such as RTs problematic (e.g. Pickett, 1998). As mentioned above, global cognitive impairments after brain lesions may occur without direct association with the insulted area, and many experimental designs may be too subtle and demanding to be used with such patients. On the contrary, the sensorimotor effects of TMS are far from compromising the ability of subjects to participate in behavioral tasks or from inducing global cognitive impairments (Walsh & Rushworth, 1999).

[64] Depending on the nature of the task, the repetition of trials and stimuli across repeated sessions for the same participant may involve learning effects that the experimenter may wish to avoid (e.g. Torriero *et al.*, 2004).

4.4. Methodological Considerations

CB TMS has been employed in a wide range of settings. Ugawa *et al.* (1991) and Amassian *et al.* (1992) first indicated that the CB may successfully undergo transcranial stimulation. CB TMS has been shown to interfere with visually guided saccades (Hashimoto & Ohtsuka, 1995), smooth pursuit eye movements (Ohtsuka & Enoki, 1998), coordinated eye and head movements (Nagel & Zangemeister, 2003), paced finger tapping (Théoret *et al.*, 2001), procedural learning (Torriero *et al.* 2004, 2007), state estimation of arm movements (Miall & Christensen, 2004; Miall *et al.*, 2007; Miall & King, 2008), and temporal processing of millisecond time intervals (Koch *et al.*, 2007). CB TMS has also indirectly modulated the excitability of the contralateral primary motor cortex (Oliveri *et al.*, 2005; Fierro *et al.*, 2007a; Koch *et al.*, 2008). It has moreover been exploratorily employed with therapeutic perspectives on patients affected by refractory epilepsy (Brighina *et al.*, 2006), and Parkinson's disease (Koch *et al.*, 2009).

Space restrictions do not allow for a review of CB- or language-related TMS studies. Work on the former has been recently conducted by Oliveri *et al.* (2007), while for the latter Devlin and Watkins (2007) provide a recent review. It is worth pointing out, however, that the studies reported here are the first linguistic TMS studies on the CB. The only two studies of CB TMS of some relevance would be Desmond *et al.* (2005) and Ferrucci *et al.* (2008); both reported disruptions of performance in the Sternberg verbal working memory task (Sternberg, 1966) after stimulation of lateral CB loci. Apparently, the level of language processing in such tasks (encoding and retrieval of letter sequences) is quite low with respect to the richness of properties of linguistic input and output. It also means that the design of the CB TMS linguistic experiments reported here, to which I now turn, had to be made in the absence of any pre-established set-ups. Decisions as those on the size and shape of the stimulating coil, on the localization method of the site of stimulation, on the choice of suitable control sites, on

the frequency, intensity, and duration of stimulation are common to all TMS studies, regardless of the content of the experimental hypotheses, and have a direct impact on the extent and direction of the behavioral changes induced.

4.4.1. Control Conditions

In general, TMS represents a rather painless method of stimulating brain loci through the scalp (section 4.2). However, it still produces a wide range of sensory inputs that may interfere with task performance. Discharge of the stimulator through the coil produces a loud click which is very difficult to conceal. Moreover, there is a definite tactile sensation on the scalp that may be due to stimulation of cutaneous nerves under the coil; stimulation of motor nerves in the scalp also induces muscle twitches. Especially in areas with large masses of muscle nearby, the sensation produced may be quite strong and even unpleasant, a fortiori with trains of high frequency stimuli, as those used here (section 4.4.5 below). Facial and trigeminal nerves may be also activated, while stimulation at any scalp site may also produce a blink reflex. In the region overlying the CB in particular, a range of adjacent muscles and nerves are liable to co-stimulation, making CB TMS far from devoid of sensation: lateral CB TMS may directly stimulate neck muscles, the brachial plexus, muscles in the neck or shoulder, and is sufficiently loud that it can provide a startling stimulus affecting speed of movement onset (Miall *et al.*, 2007).

The experimental design should apparently exclude such phenomena as holding any explanatory significance for changes in the dependent measures (Walsh & Rushworth, 1999). Three basic designs can control for such effects. A ‘‘control time’’, a ‘‘control site’’, and a ‘‘control task’’ design (Jahanshahi & Rothwell, 2000). In a control site design, the effects of TMS at a target and a control site are compared, with TMS predicted to induce an effect on the target, but not on the control site. In the control task design, the effects of TMS on experimental and control tasks are compared, with TMS

predicted to induce a performance effect on the target task employing the process of interest, but not on the control task. In a control time design, the effects of TMS are compared at several points in time during task performance, with TMS predicted to induce an effect at particular times during task performance, and not at other times. Unlike the control site and control task designs, a control time design presupposes an online rTMS design (section 4.4.3 below). Another control condition frequently employed in TMS studies is 'sham TMS', in which all parameters of TMS are maintained, with the exception that no effective magnetic field is directed toward the brain of the participant, who is not aware of the difference.

In the experiments performed here, the basic parameter setting that minimized the nonspecific effects of CB TMS on performance is the offline/distal protocol preferred (section 4.4.3). Stimulation occurred in a discrete event after the first phase of the task, and before the second. Thus, the sensory, nonspecific effects of the machine output were temporally far from influencing performance in the phase after stimulation. All experiments also employed a design combining a control site with a control stimulus type. In the first and second studies (chapters 5 and 6), participants completed the task twice in two different sessions. After the first half of each session, participants underwent TMS, and, after the end of the TMS, they completed the second half of each session. Instead of a temporally distinct control task, participants completed only one task, where their performance in two different phenomena was assessed, one of which was expected to remain unchanged after TMS (see section 4.4.11 for details). As the experiments applied a distal TMS protocol, control time conditions were a de facto unavailable option. Instead, performance before stimulation provided baseline conditions for the performance after TMS. In the third study reported (chapter 7), for task-particular reasons, participants completed only one session and underwent TMS in one of the two CB sites stimulated (section 4.4.8). A 'no TMS' group of subjects was included as a third group, that participated in the experiments in exactly the same settings but without TMS. Thus, control sites were included in the third study as a between-subjects factor (e.g. Miall & Christensen, 2004; Torriero *et al.*, 2004).

Furthermore, sham TMS was not employed, primarily because a sham coil was not always available in the laboratories. Moreover, sham TMS is often unsuccessful, especially for experienced participants, who are familiar with the sensation accompanying real TMS (G. Pobric, personal communication, 15 March 2009). Further control conditions were added by using different stimulus types (see section 4.4.11 below).

4.4.2. rTMS Protocol

In setting up the TMS experiments, it was also necessary to decide between two basic types of TMS protocols: single/ paired pulse TMS, and repetitive TMS (henceforth rTMS).

In single-pulse TMS, the effect is similar to stimulating a peripheral nerve with a conventional electric stimulator. Repeated pulses are not applied faster than once every few seconds. The pulse causes a population of neurons to depolarize and discharge an action potential, evoking measurable effects. Application of single pulses at variable times during task execution allows the investigation of the time point at which the neural activity of the stimulation site is critical for successful performance. For example, if used in the primary motor cortex, single-pulse TMS produces a motor-evoked potential (MEP), recordable on electromyography. If used on the occipital cortex, 'phosphenes' (flashes of light) might be detected by the subject. In most other areas of the cortex, the participant does not consciously experience any effect, but behavioral changes may be manifested in greater RTs, and changes in brain activity may be detected with PET or fMRI. These effects do not outlast the period of stimulation.

While initial applications of TMS involved delivery of single magnetic pulses with an approximate duration of 1 ms every few seconds, technical advancements have allowed the application of rTMS protocols, involving trains of TMS pulses at different frequencies and durations over a selected brain region (Wassermann, 1998). The higher the stimulation frequency and intensity, the greater is the disruption of brain function during the train of stimulation. Apart from the immediate effects during the train itself, rTMS modulates the excitability of the stimulated area, even beyond the duration of its application (Chen *et al.*, 1997; Pascual-Leone *et al.*, 1998). Depending on other parameters, it is possible to potentiate or depress cortical excitability (Pascual-Leone *et al.*, 1998). The general pattern is that low-frequency (1-2 Hz) rTMS reduces motor cortical excitability, while high-frequency (> 5 Hz) rTMS may increase excitability; the latter case, however, is remarkably less robust (see Walsh & Pascual-Leone, 2002).[65] Long-term potentiation and long-term depression (LTD) of cortical synapses or closely related mechanisms have been suggested as candidate explanations for the effect of high- and low-frequency rTMS, respectively (Chen *et al.*, 1997). The experiments here all employed a rTMS protocol.

4.4.3. Distal rTMS Protocol

The experimental design should also cater for setting the parameter of the TMS protocol by which control trials are distinguished from test trials. In 'online rTMS', subjects perform the task and, at a specific time-point of certain trials (usually in the order of 1-200ms), a TMS pulse is administered. This is expected to affect task-specific performance selectively in the TMS trials. In 'offline/ distal rTMS', performance at a task before rTMS (applied for a few seconds or minutes) is compared with performance after it. The apparent advantage of the online TMS protocol is its significance in identifying the time in which a particular area of brain is active in a task- the rationale being that, if the TMS pulse is administered to an area that is active at the time and is

[65] Theta-burst stimulation (section 4.4.5) is one of the many exceptions.

processing relevant information, then performance should be affected. However, a disadvantage would be the very interference of the somatosensory/ auditory input accompanying TMS and its nonspecific effects on performance (section 4.4.1 above).

The experiments here used an offline/distal rTMS protocol: the modulatory effect of offline rTMS on a brain area beyond the duration of the rTMS train allows the study of TMS-induced changes without any non-specific disruptions by TMS during task performance. In this design, behavior is evaluated before and after rTMS, rather than during rTMS. Furthermore, the opportunity to assess the particular timescale of brain processes with online TMS becomes a disadvantage where no such hypothesis is possible – a fortiori when the involvement of the area investigated is still a demonstrandum, as in the case for CB language processing.

4.4.4. Shape, Size, and Orientation of the Coil

The focus and, partially, the depth of the magnetic field depend on the shape of the stimulation coil. Two different shapes of coils are most commonly used—a circular coil and a figure-of-eight ('butterfly-shaped') coil. The former belongs to the first generation of coils and is quite powerful, inducing a non-focal, more widely distributed electric field. The latter involves more focal stimulation, with the field under the coil producing maximal current at the intersection of the two round components, where the currents flow in the same direction, converge and summate (e.g. Roth *et al.*, 1991). For CB stimulation, however, the choice is most of the time limited between a figure-of-eight coil, and a 'double cone coil', i.e. a figure-of-eight coil with the two components at an angle, and thus with increased power at the intersection.

CB stimulation has been originally performed with a double cone coil (Ugawa *et al.*, 1995), which is admittedly more adequate for deeper stimulation (Werhahn *et al.*, 1996) and thus often preferred (e.g. Ugawa *et al.*, 1995; Daskalakis *et al.*, 2004; Miall *et*

al., 2007; Miall & King, 2008). However, most CB TMS studies have successfully used the figure-of-eight coil (e.g. Hashimoto & Ohtsuka, 1995; Miall & Christensen, 2004; Oliveri *et al.*, 2005; Torriero *et al.*, 2004, 2007; Fierro *et al.*, 2007a, b; Koch *et al.*, 2007, 2008, 2009).

The experiments conducted here were constrained by the unavailability of a double cone coil in the laboratories where they were performed. Use of this coil is also marked by certain disadvantages, one being the elevated discomfort that participants often experience (N. Muggleton, personal communication, 4 May 2010). Moreover, for CB TMS, the wings tend to activate the ear and lateral neck muscles and for some head shapes, it is impossible to get the center of the coil adjacent to the scalp (C. Miall, personal communication, 16 April 2009). For instance, in the 90-mm radius double cone coil centered 3 cm lateral and 1 cm below the inion that was used by Miall *et al.* (2007), one wing of the coil overlapped the participant's right ear. Given a fortiori that the lateral site of stimulation used here is more lateral than 3 cm to the right from the inion (see sections 5.4.3 and 7.4.3), its adjacency to the right ear and face could have become a greater concern.

On the other hand, it has been further suggested that choosing a flat figure-of-eight coil over a double cone coil increases the probability that CB rTMS will be confined to the outermost CB cortical layers (Brighina *et al.*, 2006); this could be translated as a methodological advantage in favor of using the former. Behavioral changes induced could be explained with added certainty on the basis of the modulation of the excitability of the Purkinje cell layer of the CB cortex and the increased/decreased inhibitory input that it provides to the deep CB nuclei (section 4.4.9).

With respect to its size, the figure-of-eight coil used in all experiments had a diameter of 70 mm on each loop. Coils of this size have been used in most CB TMS studies (e.g. Torriero *et al.*, 2004, 2007; Koch *et al.*, 2007, 2009). A few studies have used coils of a diameter of 50 mm (e.g. Hashimoto & Ohtsuka, 1995). The broader

magnetic field generated by the larger coil was primarily motivated by the unavailability of any known CB coordinates relevant with the tasks employed here; smaller diameter coils provide more focal stimulation, which is bound to precise knowledge of brain coordinates for stimulation. Thus, selecting the latter would have constituted a methodologically unnecessary commitment. On the contrary, the larger spread of activation that a larger coil would induce in the CB cortex was considered to provide an advantage, in view of the depth of the targeted areas (M. Oliveri, personal communication, 15 May 2009). Moreover, larger coils generally have more depth of penetration than smaller ones (The Magstim Company Ltd, 2009). Thus, the size of the coil also catered for additional power in reaching the rather deep CB loci stimulated.

The coil was positioned tangentially to the scalp, with the handle pointing superiorly. The current in the coil was thus directed upward, which induced a downward current in the CB cortex. This coil position has been found to be optimal for suppressing the contralateral motor cortex in single-pulse TMS investigations (Hashimoto & Ohtsuka, 1995; Ugawa *et al.*, 1995; Daskalakis *et al.*, 2004; Oliveri *et al.*, 2005) and to interfere with cognitive processes, such as procedural learning and millisecond timing in 1 Hz rTMS paradigms (Torriero *et al.*, 2004, 2007; Koch *et al.*, 2007).

4.4.5. TMS Frequency

The frequency in which the magnetic stimuli are delivered critically determines the effects of TMS on the targeted region of the brain, both qualitatively (facilitation or disruption), and quantitatively (size of effect on dependent measure).

The rTMS protocol used here was ‘continuous theta-burst stimulation’ (henceforth cTBS). All experiments employed in particular the ‘offline’ TBS procedure (section 4.4.3 above), which has recently been applied in physiological (e.g. Di Lazzaro *et al.*, 2005) and cognitive studies (e.g. Vallesi *et al.*, 2007). In particular, for the first

and third TMS studies reported here (chapters 5 and 7), a Magstim 'Super Rapid^2' stimulator (The Magstim Company Ltd, Whitland, Dyfed, United Kingdom) was available, and thus the standard cTBS protocol (Huang *et al.*, 2005) was applied. This protocol employs a brief burst of 3 low intensity, high frequency (50 Hz, i.e. 20 ms among each stimulus) TMS pulses every 200 ms, repeated five times per second, i.e. delivered at a 5 Hz rhythm (at 'theta-frequency', in electroencephalographic terminology). On the contrary, the second study used a modified cTBS protocol (Nyffeler *et al.*, 2006), since the 'Magstim Standard Rapid^2' that was available could not support the Huang *et al.* (2005) protocol (see section 6.2.3 for more details).

In only two years after its publication, ''cTBS has become the preferred method for inducing ''virtual brain lesions'' […] to study brain-behavior relationships in motor and cognitive studies'' (Gentner *et al.*, 2008, p. 2046). It is particularly attractive because even short periods (20-40 seconds) of low-intensity TBS result in changes in cortical excitability with larger effect sizes and smaller interindividual variability than those in rTMS protocols (Huang *et al.*, 2005). In particular, when applied continuously (cTBS), it depresses cortical excitability, whereas when given in an intermittent form, i.e. 'i(ntermittent) TBS',[66] excitability is increased (Huang *et al.*, 2005). CB cTBS has only recently been applied in two studies, inducing changes in the excitability of the contralateral primary motor cortex in healthy subjects (Koch *et al.*, 2008) and, therapeutically, in Parkinsonian patients (Koch *et al.*, 2009; see section 4.4.9 below for discussion).

[66] In iTBS, a train of bursts is applied over 2 seconds, and then repeated every 10 seconds for 190 seconds.

4.4.6. Stimulus Intensity

Stimulation needs to be potent enough to induce a physiological response, resulting in effective interference with processing. The operator of TMS can control the intensity of the stimuli by changing the amplitude of current flowing in the coil, thus alternating the magnitude of the induced magnetic field and of the secondary electrical field induced. However, this involves quite an arbitrary decision, given the lack of a systematic method for determining an appropriate intensity (see Devlin & Watkins, 2007 for discussion). Some studies define the intensity on the basis of individual subjects' motor thresholds, and in particular, either the resting motor threshold, or the active motor threshold.[67] There is, however, no systematic relationship between the threshold needed to evoke an MEP and the threshold needed to evoke phosphenes in visual cortical TMS, suggesting that, outside of the motor system, motor thresholds may not be particularly appropriate (Stewart *et al.*, 2001).

A fortiori for a subcortical and even more distant locus like the CB, motor thresholds should be seen as even less significant, despite the fact that the majority of CB rTMS studies arbitrarily define intensities on the basis of motor thresholds. Here, given that most studies choose a stimulation intensity that appears to work (Devlin & Watkins, 2007), stimulus intensity was kept steady across participants to a 45% of maximum machine output; this also helped economize on the time spent for each experimental session. It is, of course, instructive to consider whether the intensity levels employed here are comparable with those used in equivalent TMS studies reporting reliable effects. In cTBS, 40% machine output is often effectively used as a fixed inter-subject stimulus intensity level, roughly corresponding to 80% of active motor threshold (N. Muggleton, personal communication, 25 August 2009). Indeed, most cTBS studies use an 80% (e.g. Huang *et al.*, 2005; di Lazzaro *et al.*, 2005, Vallesi *et al.*, 2007) or 70%

[67] The resting motor threshold is typically defined as the lowest intensity that produces MEPs of >50 lV in at least five out of 10 trials with the muscles relaxed (Rossini *et al.*, 1994). The active motor threshold is defined as the lowest intensity that produced MEPs of >200 lV in at least five out of 10 trials when the subject made a 10% of maximum contraction using visual feedback (Rothwell, 1997).

of active motor threshold stimulus intensity (Todd *et al.*, 2009). The first CB cTBS studies reported in the literature have also used the standard 80% of active motor threshold (Koch *et al.*, 2008, 2009). This means that the 45% machine output used here roughly corresponds to a 90% of a typical mean active motor threshold. The intensity used can also be roughly calculated with respect to average resting motor thresholds reported. Miall and Christensen (2004), for example, have used a mean stimulation level of 58% of maximum machine output for a 1 Hz rTMS protocol, corresponding approximately to 120% of subjects' resting motor threshold. Thus, 100% of the mean resting motor threshold corresponded to a 48% machine output, and 45% of machine output roughly corresponds to an intensity of over 90% of resting motor threshold. Indeed, such stimulation levels have been successfully used even for 1 Hz offline CB rTMS (e.g. Torriero *et al.*, 2004), which employs much higher stimulation intensities than cTBS protocols altogether. This increased machine output was used here to compensate for the depth of the areas stimulated (C. Miall, personal communication, 16 April 2009). Using stronger machine output was not considered, as 45% of machine output for 40 seconds was tolerable, yet not devoid of muscle twitches and/or irritation (section 4.4.12).

4.4.7. Duration of Stimulation

In the first and third TMS studies reported here, cTBS lasted 40 seconds. The majority of cTBS studies use a stimulation duration of 20 seconds (e.g. Di Lazzaro *et al.*, 2005; Vallesi *et al.*, 2007), whereas those using 40 seconds are significantly fewer (e.g. Todd *et al.*, 2009).The increased duration chosen here translates into double the number of TBS pulses delivered, i.e. 600 pulses instead of the 300 pulses for 20 seconds. Encouragingly, the first CB cTBS studies reported in the literature have also used 40 seconds of stimulation (Koch *et al.*, 2008, 2009).

With respect to the window of opportunity provided by cTBS, 20 seconds of stimulation are known to reduce the amplitude of MEPs for 20-30 minutes (Di Lazzaro *et al.*, 2005), and, in general, to reduce excitability for up to 20 minutes (Huang *et al.*, 2005; Vallesi *et al.*, 2007). Given that the experiments reported here involved a post-TMS phase that lasted no more than 20 minutes, the window of opportunity provided by 40 seconds of TBS was more than sufficient. However, an increase in the number of pulses has induced no manifested difference in the size of the TMS effect (Y.- Z. Huang, personal communication, 10 March 2010). Still, though, the longer period of stimulation was selected, in order to establish a longer (flatter) effect peak, with TMS effects fading out well after the session.

Moreover, in all experiments reported here, the rest period between the end of cTBS intervention and the beginning of the post-TMS block was set to 5 minutes. This was in view of the fact that reduction of cortical excitability of the motor cortex reaches a maximum from 7 to 14 min after TBS (Di Lazzaro *et al.*, 2005; Huang *et al.*, 2005). Certain cTBS studies set the rest period between TBS intervention and the post-TMS block to 10 minutes (e.g. Voss *et al.*, 2008), while other ones often do not report such rest periods (e.g. Vallesi *et al.*, 2007). The experiments here followed common practice in allowing for a 5-minute rest period after TMS (N. Muggleton, personal communication, 25 August 2009).

4.4.8. Sites of Stimulation

All experiments reported here involve contrastive stimulation of the two basic NCB sites: a right medial NCB compartment, i.e. the right superior posterior CB vermis, also known as the right 'oculomotor vermis', or 'neocerebellar vermis' (NCBV), and a right lateral NCB compartment. Contrastive activation of the lateral CB and vermis has been performed before in a paced finger-tapping task (Théoret *et al.*, 2001). The two

sites belong to the two broader areas found to be activated in linguistic tasks in imaging studies, and are marked with the red elliptical circles in figure 4.1 below. An apparent disadvantage for the TMS experiments reported here is that the two sites were defined in the absence of any known CB activations particular to the tasks used (see section 8.2.1).

Anatomically, the oculomotor vermis, which belongs to the superior posterior vermis, is one of the compartments closest to the TMS coil (Miall & Christensen, 2004), and its stimulation has induced behavioral effects with high spatial precision (Hashimoto & Ohtsuka, 1995). On the contrary, the depth of PLCB loci often makes the use of focal coils an issue (section 4.4.4 above). This might explain why in the first and second study TMS of the lateral site provided a satisfactory control condition (sections 5.4.3 and 6.2.3). On the contrary, in the third study, a different lateral site was stimulated, with smaller depth from the scalp surface, yielding effects after TMS (section 7.4.3). Regarding the retrieval of the sites for stimulation, previous investigations have demonstrated that CB TMS predominantly affects the posterior and superior lobules (Hashimoto & Ohtsuka, 1995; Ugawa *et al.*, 1995). This is encouraging, given that the areas of interest are found in the superior posterior CB compartments, as shown in figure 4.1 below.

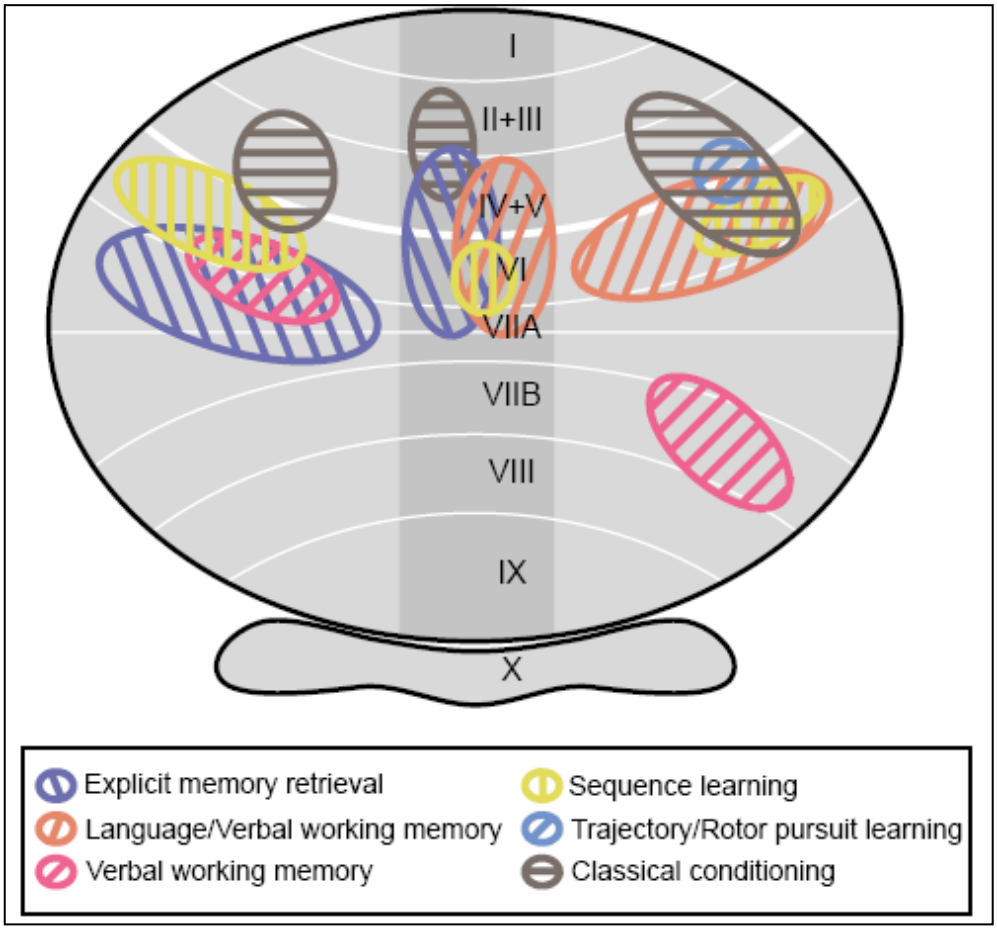


Figure 4.1: Schematic diagram of the unfolded CB with activation sites in the review of Desmond & Fiez (1998). Roman numerals appearing in the cerebellar vermis denote the lobule identification based on Larsell & Jansen (1972). The line separating IV+V from VI represents the primary fissure. The two red elliptical circles in the right CB are also the two NCB language-related sites that were stimulated in the experiments reported here. Figure reproduced from Desmond & Fiez (1998, p. 356) with permission © 1998 Elsevier.

With respect to successful targeting of the TMS sites, accurate positioning of the stimulating coil is achieved on the basis of functional, anatomical, or heuristic criteria and methods (Devlin & Watkins, 2007).

One approach is to identify the stimulation site using a 'functional localizer' task over various sites within a predefined region. Upon observing the expected TMS-induced effect, the site is marked and used for subsequent stimulation. The apparent disadvantage with this approach is the very absence of a quick-and-easy task applicable for identifying a relevant locus for CB linguistic function.

Alternatively, functional imaging can be used to identify activated brain regions in participants which can be subsequently targeted with TMS. Interestingly, however, Desmond *et al.* (2005) represents the only known CB TMS study that has utilized this method. While establishing accuracy in stimulation, this approach is considerably demanding with respect to the available resources (time required, fMRI availability, staff with technical expertise in operating the fMRI equipment).

Another method uses head-surface digitization and registration of the stimulation sites onto three-dimensional reconstructed head MRIs. In that way, the actual brain target is individually specified for each experimental subject. While this method provides high spatial resolution, availability of MRI images for participants would translate into a further exclusion criterion on the subject pool – the requirement for the participants to be native speakers of English was already quite constraining. Another reason was the lack of any coordinates for CB activations in the imaging literature for the tasks employed here.

Finally, the 'heuristic approach' employs population-based estimates of the underlying anatomy, where TMS is applied in relation to specific individual scalp landmarks, such as the vertex or the inion. This approach is certainly disadvantaged in different ways; for example, it does not account for interindividual differences in the correspondence between scalp landmarks and underlying brain anatomy; furthermore, it is blind regarding any interindividual differences in the functional organization of the brain. Thus, this approach is bound to introduce errors in the targeted brain region (Meyer *et al.*, 1991). Procedurally, however, the heuristic approach is arguably the most

flexible one, as it does not require available brain images for participation or neuronavigational software. As such, the heuristic approach was preferred here.[68] Indeed, the majority of CB TMS studies do not use a neuro-navigation system (e.g. Hashimoto & Ohtsuka, 1995; Ohtsuka & Enoki, 1998; Daskalakis *et al.*, 2004; Miall & Christensen, 2004; Koch *et al.*, 2007; Miall *et al.*, 2007; Miall & King, 2008). Instead, they often rely on suggestive evidence of scalp-brain correspondences yielded from a small number of pilot subjects that undergo head surface digitization and registration (e.g. Brighina *et al.*, 2006; Del Olmo *et al.*, 2007; Koch *et al.*, 2007). Similarly, here, the registered MRI image of a volunteer was recruited in order to convert any relevant Talairach coordinates used into scalp-based coordinates, or reversely, to calculate approximate Talairach coordinates from scalp-based ones (see sections 5.4.3 and 7.4.3). For all transformations, the Brainsight™ TMS-MRI co-registration system (Rogue Research, Montreal, Canada) was used.[69] For higher precision, stimulation sites were marked with non-permanent colored markers on participants' scalps, instead of the often-used lycra caps where stickers are attached.

4.4.9. Cerebellar Mechanism Affected

I now turn to considerations on the changes that cTBS might induce on CB microcircuitry. Any direct effect of rTMS on the deep CB nuclei is considered implausible: especially when using a focal coil, the effects are limited to the first cortical layers of the stimulated area (Oliveri *et al.*, 2005, 2007), and do not go beyond 2-3 cm below the scalp[70] (e.g. Fierro *et al.*, 2007b). TMS effects on the CB are primarily explained on the basis of the potentiation or depression of Purkinje cell activity. For example, single-pulse TMS on the CB has been used to condition the excitability of motor cortex; this is consistent with the brief excitation of Purkinje cells that induces a

[68] See section 8.2.1 for further discussion.
[69] Thanks to Dr. Neil Muggleton for volunteering and helping in the registration process.
[70] This seems to be an overestimation; the depth is usually considered to be around 1.5-2 cm (see section 4.2).

brief silencing of CB nuclei, leading to CB inhibition over the motor cortex via cerebello-thalamo-cortical pathways (Ugawa *et al.*, 1995; Werhahn *et al.*, 1996).

With respect to the particular TMS protocols used here, direct cTBS on the motor cortex has been found to temporarily depress cortical activity, similarly to 1 Hz rTMS (Di Lazzaro *et al.*, 2005; Huang *et al.*, 2005; Silvanto *et al.*, 2007; Vallesi *et al.*, 2007). In the same vein, 1 Hz rTMS on the CB has been proposed to disrupt activity of the inhibitory Purkinje cells, thus reducing the drive from the dentate and interposed nuclei to the cortex via the thalamus (Daskalakis *et al.*, 2004). For example, 1Hz rTMS over the left CB cortex has been found to increase intracortical facilitation of the right motor cortex, with low-frequency rTMS conceivably decreasing Purkinje cell inhibition on deep CB nuclei, thus ultimately enhancing thalamo-cortical facilitation of cortical outputs (Oliveri *et al.*, 2005).

Based on the similar effects of 1 Hz rTMS and cTBS, it would be assumed that temporary reduction of Purkinje cell excitability should result in reduced inhibition of the deep CB nuclei, ultimately leading to increased excitation of their cortical outputs via the thalamus. However, the first results of cTBS on the CB have been in the very opposite direction: MEP amplitude obtained from the contralateral primary motor cortex decreased after PLCB cTBS. This unexpected finding has been speculatively explained on the basis of plastic changes induced in both excitatory and inhibitory synapses of the cerebello-thalamic and thalamo-cortical pathways, involving sub-populations of interneurons with lower thresholds of excitability. Such effects would eventually lead to different changes in the excitability of the primary motor cortex, as compared with those induced by 1 Hz rTMS (Koch *et al.*, 2008).

According to the NCBKFLP hypothesis here, the Purkinje cell layer receives a copy of the state of the language-related cortical loci and regulates, via inhibition, the output of the deep CB nuclei; the latter will transmit, via the thalamus, the linguistic prediction of the NCB internal models back to the cortex (section 2.3.3). Thus, CB TMS

is expected to affect this process of linguistic emulation by modulating the excitability of the Purkinje cells of the CB cortex and thus indirectly interfering with the deep CB nuclear output to the cortex.

However, as seen above, the first CB cTBS study has unexpectedly yielded effects in the opposite direction of those with 1 Hz rTMS. This means that the direction of the effects induced in the studies reported here cannot be explained with much certainty. Rather, the primary emphasis will be placed on the selective nature of the effects (for some linguistic stimuli and not others; see section 4.4.11 below).

At an elementary level, though, an idea that will be entertained below is that of performance enhancement due to disrupted Purkinje cell inhibitory functions. Enhancement in cognitive performance is best explained on the grounds of neural noise introduced into an inhibitory component of the processing system, rather than as potentiation in processes directly contributing to the output of the task (Walsh & Pascual-Leone, 2003, p. 93; section 4.4.2 here). As seen above, the Purkinje cell layer is the fundamental inhibitory component that is accessible to the TMS coil. Thus, the enhancements found in linguistic performance in the subsequent chapters will be briefly discussed in the light of a possible disruption of Purkinje cell inhibition to the constantly excited CB nuclei.

4.4.10. Dependent Measures

Regarding the selection of the appropriate dependent measure, ‘‘the choice of dependent variable in a TMS experiment depends on the function to be disrupted, but reaction time is proving to be a more versatile dependent measure than error rates’’ (Walsh & Pascual-Leone, 2003, p. 65). In TMS language studies in particular, the primary dependent measure is bound by the very nature of the paradigm to RTs. In contrast to patient studies, which typically measure behavioral deficits in terms of reduced accuracies, the virtual lesions induced by TMS generally become manifest as changes in RTs. This would be because TMS does not inactivate a region in the same way that a lesion does; instead, it introduces ‘neural noise’ into the computation, leading to RT increases or decreases, rather than error rates. Conceivably, the spared information in the neural circuit suffices to overcome the noise by means of additional processing time, hence the increased RTs (Devlin & Watkins, 2007). On the other hand, decreased RTs would most probably relate with explanations of disrupted inhibitory processes (section 4.4.9, above).

4.4.11. Priming Paradigm

However, the employment of RTs as the primary dependent measure in TMS paradigms imposes quite directly here a constraint on the employable tasks. Many phenomena that would be of relevance for the central hypothesis are in fact captured on the basis of accuracy rates as the primary dependent measure (section 2.6). For example, Ferreira’s (2003) misinterpretation of non-canonical sentences or semantic illusion studies (Erickson & Mattson, 1981) would not be the optimal choice for conducting a CB linguistic TMS experiment.

Furthermore, an experimental paradigm with sentential stimuli would introduce a wide range of factors that could be affected by TMS, but would not be directly relevant with the neurolinguistic hypothesis (chapter 2), e.g. working memory or morphosyntactic complexity. Given that the experiments reported here were the first linguistic CB TMS studies in the literature, it was necessary to narrow down the aspects that TMS would conceivably influence. These constraints motivated the adoption of a lexical priming paradigm, as studied in a standard lexical decision task (Meyer & Schvaneveldt, 1971). Lexical priming paradigms may yield findings with far-reaching consequences for processing at the sentence level, a fortiori beyond the constraints involved in the latter:

> the priming task […] represents a lower bound on the information available to the comprehender by eliminating constraints offered by the other nouns in the sentence, case marking, context, and so on. Word-word priming, as a measure of what is activated in the absence of other constraints, thus offers a stringent test for studying these phenomena (McRae *et al.*, 2005, p. 1176).

Lexical decision tasks involve the presentation of letter strings, on which a subject is required to decide whether they form a word or not (by button presses or oral decisions). The presentation of a related word (the prime) prior to the target string of letters makes the decision faster and more accurate relative to the previous presentation of an unrelated word (e.g. Cañas & Bajo, 1994). The lexical decision paradigm offers the possibility of establishing a dissociation within a unitary task between two or more different cases of priming, instead of introducing a distinct control task, as is often done in TMS research (section 4.4.1). In all studies reported here, the control stimulus type involved categorically related pairs (e.g. ‘penny-coin’), while the experimental stimulus type involved associatively related pairs (e.g. ‘gift-horse’). Thus, the possible issues that

might arise for the addition of another task, such as counterbalancing the order of the two tasks, were simply avoided here.

Following a standard convention in neuropsychological research employing priming paradigms, the dependent measure was the difference between the mean RTs for the unrelated items minus the mean RTs for the related items per condition (e.g. Ober, 2002; Rogers & Friedman, 2008). The same calculation was performed over accuracy rates. In each case, the ratios between the two means were also calculated; analyses based on ratios did not compromise the significance or the effect sizes reported. For the first and third TMS studies, the results of pilot experiments are also presented, showing significantly shorter latencies for related as compared to unrelated pairs, thus confirming that priming occurred. In the second study, the number of participants of the main study sufficed to show that priming occurred directly on the basis of such latencies.

In all experiments reported, lexical decision latencies longer than 3 SD from the mean reaction time for correct 'word' responses (right-hand index finger responses) were excluded; the same trimming was applied for the correct 'nonword' responses (left-hand index finger responses). Word- and nonword-latencies were trimmed separately, given that TMS on the right CB could selectively affect right-hand index finger control via its projections to the contralateral (left) motor cortex, which in turn controls its contralateral (right) limbs. Latencies longer than 1200 ms were also excluded beforehand, as reflecting low familiarity with the stimulus or distraction, rather than lexical access (e.g. Perea & Gotor, 1997). Including the trials that received latencies longer than 1200 before applying the 3 SD cut-off point in trimming did not compromise the size or the significance of the effects reported in these analyses in any of the experiments. Accuracy rates per condition were arcsine transformed and submitted to a separate ANOVA (e.g. Moss *et al.*, 1995). Finally, as in similar experiments (e.g. McRae *et al.*, 2005), the lists across which participants and items were rotated were included as between-subjects and -items dummy variables, in order to stabilize variance resulting from such rotation (Pollatsek & Well, 1995). Analyses performed without

those dummy variables did not compromise the size or significance of the effects reported overall, unless otherwise stated. Also, wherever applicable, the Greenhouse-Geisser, Huynh-Feldt, and lower-bound corrections for violations of sphericity did not yield considerably different effect sizes or significance levels.

4.4.12. Subject Selection Criteria

In the experiments reported, a number of subject selection criteria were applied. These were dictated both by TMS-related safety issues, and by the nature of the tasks involved.

During a TMS session, subjects may experience headaches or nausea, or may find the face twitches and other peripheral effects of TMS considerably uncomfortable; in such case, subjects are released from any obligation to continue in the experimental session. Above all, unlike single-pulse TMS, rTMS carries a serious risk of epileptic seizures (Pascual- Leone *et al.*, 1993). In fact, seizures may even be induced in subjects not associated with any risk factors (Wassermann, 1998; Anand & Hotson, 2002). Adherence to safety guidelines was thus crucially important, and subjects falling under any of the categories below were not allowed to participate.[71]

Pregnant or breastfeeding women, or women not using a reliable method of birth control; subjects with a history of epileptic seizures, both for themselves and for their first degree relatives; subjects with a history of neurosurgery or head injury; subjects with a history of chronic medical conditions of any sort; subjects with a history of migraines or frequent headaches; subjects with a history of hearing loss and cochlear implants; subjects with a history of current hypertension, heat convulsion, or systemic and metabolic disorders; subjects with a cardiac pacemaker, intra-cardiac lines,

[71] See also Appendices E and F for the questionnaires used for the TMS on native speakers of English and Modern Greek, respectively.

implanted neurostimulators, or any other implanted electronic device (the field pulse would disturb nearby electronic devices), as well as intracranial metallic implants, surgical clips, medical pumps, or other metallic implants outside the mouth, or metallic particles in the eye (the magnetic field would generate forces on objects exposed to it); subjects with a history of alcohol abuse, drug dependency, or significant psychiatric illness; subjects with inadequate communication skills or under custodial care.

Subjects were not allowed to have participated in another TMS experiment earlier on the same day. This was required both to avoid any confounding effects of stimulation of another brain area, as well as for safety reasons (N. Muggleton, personal communication, 25 August 2009). In the first and second TMS study (chapters 5 and 6), where each participant underwent TMS in two different sessions on two different sites, a minimum distance of 2 days mediated between the first and second TMS session. This was required in order to avoid confounding effects of the stimulation of the initial site on that in the second session (e.g. Walsh & Pascual-Leone, 2003).

With respect to the tasks at hand, participants were required to have normal or corrected-to-normal vision, since the tasks involved reading of letter strings. For the same reason, dyslexic subjects could not participate. Subjects were also required to have normal motor functions, since the task involved rapid button presses in response to the visual stimuli. Only right-handed participants were recruited. This constraint, often applied by default in TMS experiments (G. Pobric, personal communication, 10 March 2009), helps keep factors like hemispheric dominance for both motor and language functions steady across participants. Subjects were also asked if they had any attentional deficits. No such participants were included, as stimulus perception and response generation could be compromised. Similarly, subjects that participated in the pilot experiments before the main TMS study were all right-handed, with normal or corrected-to-normal vision, with no known attentional, motoric, or reading problems.

4.5. Conclusion

In this chapter I provided the reasons why TMS experimentation was preferred over a clinical approach. I described the particular settings of TMS that were selected and the reasons for doing so. As the following studies reported are the first TMS studies on CB language functions, their TMS setup had to be constructed de novo, in the absence of any established paradigms.

Chapter 5

TMS Study 1: Effects of Theta-Burst Stimulation of the Right Neocerebellar Vermis on Formal-Associative Priming

‘‘Tonight's the night I shall be talking about of flu the subject of word association football. This is a technique out a living much used in the practice makes perfect of psychoanalysister and brother and one that has occupied piper the majority rule of my attention squad by the right number one two three four the last five years to the memory. It is quite remarkable baker charlie how much the miller's son this so-called while you were out word association immigrants' problems influences the manner from heaven in which we sleekit cowering timrous beasties all-American speak, the famous explorer. And the really well that is surprising partner in crime is that a lot and his wife of the lions' feeding time we may be c d e effectively quite unaware of the fact or fiction section of the Watford Public Library that we are even doing it is a far, far better thing that I do now then, now then, what's going onward christian Barnard the famous hearty part of the lettuce now praise famous mental homes for loonies like me. So on the button, my contention causing all the headaches, is that unless we take into account of Monte Cristo in our thinking George the Fifth this phenomenon the other hand we shall not be able satisfact or fiction section of the Watford Public Library againily to understand to attention when I'm talking to you and stop laughing, about human nature, man's psychological make-up some story the wife'll believe and hence the very meaning of life itselfish bastard, I'll kick him in the balls Pond Road.’’ (Cleese, 1973). [72]

[72] Many thanks to Dr. Patrick Sturt for the suggestion.

5.1. Introduction

This chapter presents the first TMS study conducted. This is, moreover, the first linguistic cerebellar (CB) TMS study in the literature, and it is also reported in Argyropoulos (2010b).

At an elementary level, a direct prediction generated by the NCBKFLP hypothesis (Argyropoulos, 2009; section 2.3.3 here) is that phenomena traditionally termed as 'lexical associative priming' should fundamentally involve NCB circuitry (section 2.3.4.3). In other words, non-attended expectancy generation for upcoming word B based on currently processed word 'A', where 'A' and 'B' are temporally contiguous in speech, should provide one of the simplest cases of NCB prediction in language processing (section 2.6.2.1).

With respect to the neuroanatomical aspects of lexical priming, research in Alzheimer's disease and semantic dementia has established that '(categorical) semantic', but not 'associative' priming is disrupted, with the latter type often providing the unaffected baseline conditions (e.g. Rogers & Friedman, 2008). However, the reverse dissociation has not been established in any clinical population; for instance, both Wernicke's and Broca's aphasia involves preserved 'semantic' and 'associative' priming (Hagoort, 1997; section 2.4.4). On the other hand, as reviewed above, there is considerable evidence supporting CB involvement in associative linguistic computations (section 2.3.4.3). Representative highlights would be the NCB deficits in sensorimotor and cognitive associative learning (Drepper *et al.*, 1999), and in generating semantic-associatively-related verbs for noun-stimuli (e.g. chef-cook, pill-take; Gebhart *et al.*, 2002; section 2.4.4.5 here). The same idea has also been entertained in Deacon's work, where the CB is thought to contribute to language ''by providing access to relatively automatic word-sequence subroutines'' (Deacon, 1997, p. 275).

However, the evidence available on NCB involvement in verb generation tasks could only support the significance of the NCB in processing semantic-associative relations; formal-associative ones, though, have not been investigated. Moreover, these findings could be misleading with respect to the nature of the deficit per se. Gebhart *et al.* (2002), for example, in the line of Thach (1996a), promote the view that the PLCB is responsible for the production of linguistic responses in a stimulus-response modality, corroborating a 'motor output'-based metaphor of CB functions (e.g. section 2.3.1 here). On the contrary, establishing that the CB computations pertain equally to production and perception in language processing as in other behaviors (sections 2.2.2 and 3.5.3) would favor a computation-based, modality-independent approach, in line with the NCBKFLP here.

Thus, in order to investigate the involvement of the NCB in processing formal-associative relations from a perception-comprehension standpoint, a lexical decision task was prepared for a TMS experiment, contrasting formal-associative with semantic-categorical priming.

5.2. Pilot Experiment 1 (Experiment 1)

In order to demonstrate that the related pairs in the task involved priming independently of any TMS effects, a pilot experiment was run first.

5.2.1. Stimuli

Each session contained 600 trials (word/nonword-word/nonword pairs). In order to avoid onset effects (Forster & Davis, 1991), primes and targets always differed in the first phoneme; for instance, a pair like 'bread-BUTTER' could not be selected. It was also ensured that primes and targets were not orthographically similar (e.g. Rogers & Friedman, 2008). This was quantified here by excluding any pairs that had a sequence of at least two letters shared. It was also ensured that the graphemic structure of the nonwords resembled that of orthographically legal words.

The session was divided in two blocks of 300 trials each. Each block consisted of two 'miniblocks' of 150 trials each, containing 60 word-word, 60 nonword-nonword, 15 nonword-word, and 15 word-nonword trials. This yielded an equal number of 'yes' and 'no' responses, as well as an equal number of words and nonwords. In the second and fourth miniblocks, the same word and nonword pairs were presented as those of the first and third miniblocks, respectively. This enabled the study of any changes in practice-related effects after CB TMS. Printed frequency was taken from the British National Corpus (BNC, written part; Leech, 1992), a 90-million-word collection of samples of written British English from a wide range of sources in the later 20th century. Prime and target words did not differ in frequency or length across blocks or miniblocks (all *Fs*, $F < 1$; table A.2 in appendix A).

The first and third miniblocks contained 15 target pairs and 15 control pairs for each of the two different priming sets-types (see table A.1 in appendix A): semantic-

categorically related pairs, where the prime word is a subordinate term of that in the target position, but does not form a phrase in speech with it (e.g. apple-FRUIT); and formal-associatively related pairs, where the prime word co-occurs in speech with but is categorically unrelated to the target word (e.g. gift-HORSE). Two lists were created, across which subjects were rotated. The stimuli of the second half of the first list were the stimuli of the first half of the second list, and vice versa (table 5.1). Changes in priming sizes were captured by comparing priming sizes between the first and third miniblocks. These were calculated according to the description in section 4.4.11. Given that pairs were not counterbalanced for relatedness, but belonged either to the unrelated or to the related items set, an items-based analysis of priming sizes was not possible.[73]

The relatedness proportion, i.e. the ratio of related per unrelated words was 0.2, as in other experiments (e.g. Rogers & Friedman, 2008). Such low proportion minimizes predictability and maintains an automatic level of priming, given that strategic components may be introduced otherwise, affecting priming sizes (e.g. Cañas & Bajo, 1994; Ober, 2002).

Half	**Priming Type**	**List**			
		A		**B**	
		Unrelated	**Related**	**Unrelated**	**Related**
First Half	**Semantic-Categorical**	apple-gem	penny-coin	pants-toy	knife-cutlery
	Formal-Associative	computer-roof	pigeon-hole	tomato-path	gift-horse
Second Half	**Semantic-Categorical**	pants-toy	knife-cutlery	apple-gem	penny-coin
	Formal-Associative	tomato-path	gift-horse	computer-roof	pigeon-hole

Table 5.1: Task conditions (experiments 1, 2, and 3).

[73] In lexical decision tasks, it is not necessary to rotate items across related and unrelated pairs-conditions; this is the case, for instance, in randomly composed lists generated for each subject (e.g. Perea & Gotor, 1997). Such a rotation would introduce additional conditions, which would be difficult to cater for with the small TMS subject pool size available. An items-based analysis was made possible only by using raw latencies as a dependent measure, entering 'relatedness' as a between-items variable. A subjects-and items-based analysis of raw latencies for the TMS study (experiment 3 here) is provided at the end of Appendix A.

As in other lexical decision tasks, target words of the formal-associative group were not matched with those of the semantic-categorical group (e.g. Ferrand & New, 2003): while the former were not less frequent (stimulus set: $F < 1$), they were shorter than the latter (stimulus set: F (1,112) = 11.62, MSe = 2.97, $p < .005$). Word frequency and length for primes and targets were thus matched across miniblocks separately for each priming type (miniblock: all *F*s, $F < 1$; tables A.3 and A.4 in Appendix A).

As often noted in the priming literature, it is difficult to separate association strength from semantic-featural overlap. This is because associative and categorical norms ‘‘are not nearly as pure as has been assumed’’ (Hutchison, 2003, p. 787). For instance, semantic similarity ratings based on behavioral normative studies are often influenced by associations of items; highly-associated items tend to share semantic relations as well, and the dominance structure of categorical norms may reflect association strength between the stimuli (*ibid*). This is often why very few studies of pure associative priming have been conducted so far. Following one of these few studies, association values and semantic similarity for both ‘‘pure semantic’’ and ‘‘pure associative’’ pairs (Ferrand & New, 2003) were controlled for.

Associative strength was determined on the basis of free word generation norms provided in the Edinburgh word Association Thesaurus (Kiss *et al.*, 1973), and by the frequency of co-occurrence of these pairs in the BNC. As the experimental hypothesis of this study pertained to formal-associative priming, it was ensured that all associated pairs co-occurred in speech in an immediate fashion. For example, a pair like ‘gift-HORSE’ was selected, but not a pair like ‘storm-TEACUP’. Formal-associatively related pairs co-occurred as phrases much more frequently in the BNC than the semantic-categorically related ones across the two miniblocks (stimulus set: F (1, 56) = 15.23, MSe = 23.82, $p < .001$; rest of *F*s, $F < 1$), and had higher word-association values (stimulus set: F (1, 56) = 165.83, MSe = 0.13, $p < .001$; rest of *F*s, $F < 1$; see table A.5

for mean values in appendix A).[74] Backward priming strength for the lexical associate set was kept quite low: the EAT response strength for the prime word, given the target word as stimulus, was always between a minimum of 0 and a maximum of 0.05.[75]

Semantic-categorically related pairs were drawn from the published category norms of McEvoy and Nelson (1982), and Van Overschelde *et al.* (2004); only a subset of these contained non-periphrastic superordinate terms (e.g. 'grammatical part of speech' vs 'rodent'). The non-periphrastic pairs selected were assessed on the grounds of their semantic similarity, which was determined using the ''WordNet::Similarity'' software (Pedersen *et al.*, 2004).[76] Semantic-categorically related pairs received much higher similarity ratings than the formal-associatively related ones across the two miniblocks (stimulus set: F (1, 56) = 263.12, MSe = 0.26, $p < .001$; rest of Fs, $F < 1$; see table A.5 in Appendix A).

[74] An apparent disadvantage for stimulus construction here was that no explicit criterion was used to distinguish between semantically and formally associated pairs. The degree of semantic association of formally associated words is a poorly explored issue (see Hutchison, 2003 for discussion). Thus, certain pairs in the set of formal associates could also involve words that are semantic-associatively related: e.g. 'swan-lake' involves an association between swans and the location where they can be found. However, all word pairs of this set were formal associates, i.e. they formed idiomatic phrases.

[75] These numbers represent the ''proportion of occurrence'' (Kiss *et al.*, 1973), which is the number of times a certain word was produced in response to a certain stimulus, divided by the number of people that were given the particular stimulus.

[76] 'WordNet::Similarity' contains similarity measures making use of different aspects of 'WordNet' (hierarchy and graph structure; Fellbaum, 1998). These measures are implemented as Perl modules, taking as input two concepts, and returning a numerical value that represents the degree to which the two concepts are similar in terms of their categorical relatedness. WordNet is a lexical database where each meaning of a word is represented by a 'synonym set'. Each synonym set has a gloss that defines the concept that it represents. For example the words 'car', 'auto', 'automobile', and 'motorcar' constitute a single set that has the gloss: 'four wheel motor vehicle, usually propelled by an internal combustion engine'. Synonym sets are connected to each other through explicit semantic relations.

5.2.2. Procedure

Prime and target stimuli were presented in green colors at the center of a black screen of a MacBook using the 'DirectRT' (Jarvis, 2008) software. Each trial consisted of a sequence of three stimuli presented at the same screen location. First, a fixation point ('+') was presented for 400 ms, followed by the prime in lower-case letters for 100 ms, which was followed by the presentation of the target word. The targets remained on the screen until participants responded. No masking mediated the presentation of the prime and the target word, and thus the SOA value was confounded with that of the prime duration (100 ms). SOAs longer than 200-250 ms are generally considered to involve strategic effects (e.g. Perea & Rosa, 2002), and thus were avoided. There was an intertrial interval of 600 ms.

Participants were instructed to focus on the fixation point, read the first letter string, and respond only to the second one. The order of stimulus presentation was randomized for each subject. Subjects were instructed to press one of two buttons on the keyboard ('j' for yes and 'f' for no) to indicate whether the target letter string was an English word or not, as rapidly and accurately as possible. They used their dominant right-hand index finger for the word responses. Each subject received a total of 20 practice trials prior to the 600 experimental trials. The whole session lasted approximately 35-40 minutes, depending on the time participants took in between miniblocks. Participants were tested individually in a silent and dimly lit room. They received written instructions explaining the task; they provided informed consent and received compensation for each session.

5.2.3. Participants

18 students of the University of Edinburgh participated in this experiment; they were recruited according to the selection criteria in section 4.4.12, and were rotated across the two stimulus lists (section 5.2.1).

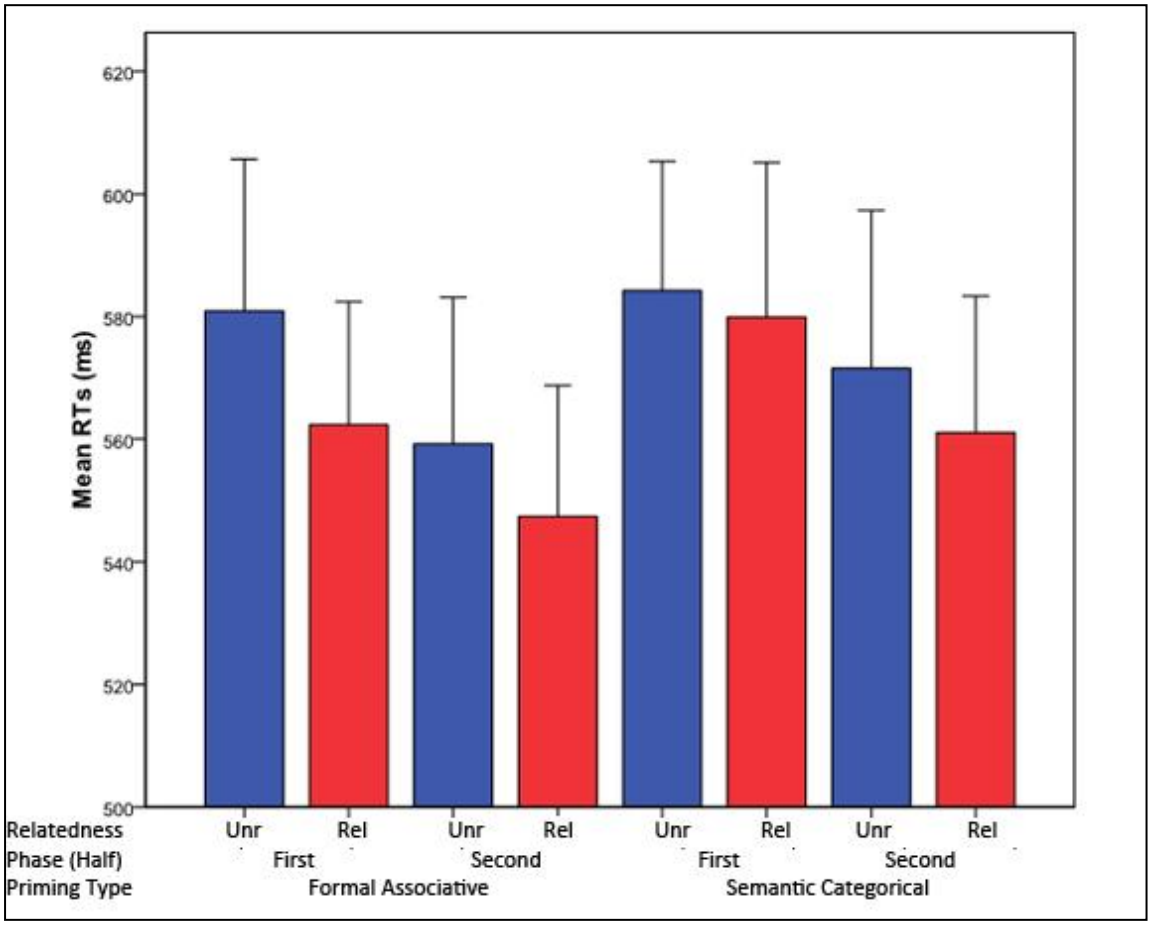
5.2.4. Design

The dependent measure here was lexical decision latencies. A 2 x 2 x 2 within-subjects design was employed: the pairs could be unrelated or related (relatedness); they belonged to the formal-associative or to the semantic-categorical group (priming type). Latencies were assessed in the first and second halves of the session (phase).

5.2.5. Results

The analysis was performed according to the specifications in section 4.4.11. Excluded trials comprised 2.4% of the data. Latencies per condition can be seen in plot 5.1 and table 5.2 below.

As demonstrated by a **three-way ANOVA (relatedness, priming type, phase)**, latencies were longer for the pairs (related or unrelated) of the semantic-categorical set, significantly by subjects and marginally by items (priming type: F_1 (1, 16) = 2.07, MSe = 975.25, $p < .05$; F_2 (1, 112) = 3.71, $p = .06$). Latencies also became significantly shorter in the second half of the session (phase: F_1 (1, 16) = 4.65, MSe = 2245.90, $p < .05$; F_2 (1, 112) = 13.68, MSe = 1248.51, $p < .001$) independently of priming set (priming type*phase: both *F*s, $F < 1$). Most importantly, **related items received shorter latencies than unrelated ones, but significantly only by subjects and marginally by items (relatedness: F_1 (1, 16) = 12.91, MSe = 356.26, $p < .005$; F_2 (1, 112) = 2.90, MSe = 8616.95, $p = .09$).** Priming occurred independently of priming type or phase in the session (all *F*s, $F < 1$, or $p > .2$).



Plot 5.1: Mean latencies for unrelated and related pairs for pilot group per condition. Error bars represent + 1 SEM (experiment 1).

Priming Type	Phase	Relatedness	Mean RT (ms)	SD (ms)
Formal-Associative	**First Half**	**Unrelated**	580.92	104.81
		Related	562.36	85.07
	Second Half	**Unrelated**	559.25	101.25
		Related	547.41	90.59
Semantic-Categorical	**First Half**	**Unrelated**	584.24	89.28
		Related	579.92	107.07
	Second Half	**Unrelated**	571.58	109.20
		Related	561.10	94.34

Table 5.2: Mean latencies for unrelated and related pairs for pilot group per condition (experiment 1).

5.2.6. Discussion

The results demonstrated priming effects independently of priming type. Such direct differences in lexical decision latencies between unrelated and related trials were necessary to show, as the main TMS experiment employed a smaller number of participants that did not independently suffice to demonstrate significant priming effects across conditions; the dependent measure there was priming size, calculated on the basis of the description in section 4.4.11. The fact that priming was only significant by subjects may suggest that only a subset of items was responsible for the priming effects, or that the number of participants was not big enough to stabilize variances for the analysis by items. However, this finding sufficed here, as in the main experiment the design was such that priming sizes could only be calculated by subjects (section 5.2.1). Importantly, priming sizes did not change in the second phase, suggesting that any changes after TMS in the main study should be interpreted as genuine effects induced by CB stimulation.

5.3. Pilot Experiment 2 (Experiment 2)

As the participants in the main TMS study were asked to complete the same task in two different sessions, a pilot group completed the two sessions in the same fashion, in order to observe any learning-related differences in the priming sizes between the first and second session that should be expected in the main TMS study.

5.3.1. Stimuli

The same stimuli were used as those in the first pilot experiment (section 5.2.1).

5.3.2. Procedure

The same procedure was used as that in the first pilot experiment (section 5.2.2).

5.3.3. Participants

8 participants were recruited according to the selection criteria in section 4.4.12, and were rotated across the two stimulus lists. Participants underwent the same protocol as those in the main TMS experiment (section 5.4) without the employment of TMS, completing two sessions on two different days. Participants also took a necessary break after the first half of each session for 7 minutes, corresponding to the time required to apply TMS and to the subsequent 5-minute break time for participants in the main TMS experiment. The two subject groups did not differ with respect to the mediating time distance between the first and second session (here: mean time distance: 14.5 days, SD: 12.39; $F < 1$). TMS participants (section 5.4.4) were not significantly older than the pilots here (mean age: 21.4; SD: 2.77 years; group: F (1, 14) = 3.95, MSe = 66.31, p = .11).

5.3.4. Design

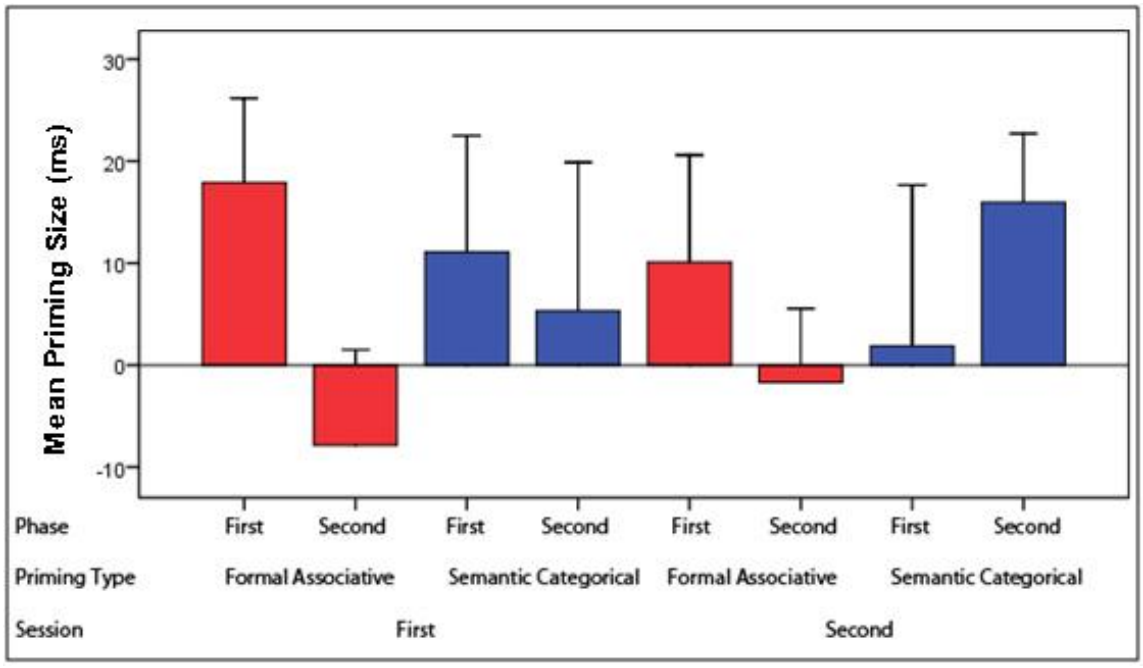
A 2 x 2 x 2 within-subjects design was employed: participants completed the task twice, in two different sessions (session number); priming sizes were assessed in the first and in the second half of each session (phase); priming sizes pertained to two different priming types: formal associative and semantic-categorical priming (priming type). Priming size was measured as the difference between the RT mean for unprimed target words minus that for primed target words (section 4.4.11).

5.3.5. Results

The trials excluded according to the criteria in section 4.4.11 amounted to 2% of the data. The mean priming sizes per condition can be seen in table 5.3 and plot 5.2 below. A three-way ANOVA (session number, priming type, phase) showed no difference in priming sizes between sessions, priming types, or phases (all *F*s, $F < 1$). There was only a marginal session number*phase interaction (F (1, 6) = 4.24, MSe = 370.97, $p = .09$) and a nonsignificant priming type*phase interaction (F (1, 6) = 3.17, MSe = 1274.25, $p = .13$; rest of *F*s, $F < 1$). These were largely due to a weak tendency for formal-associative priming size to decrease in the second phase across sessions. This was demonstrated by a two-way ANOVA for formal-associative priming sizes (phase: F (1, 6) = 3.74, MSe = 797.87, $p = .1$; rest of *p*s, $p > .25$), which was not found for semantic-categorical ones (all *p*s, $p > .35$).

Session	Priming Type	Phase (Half)	Mean Priming Size (ms)	SD (ms)
First	Formal-Associative	First	17.91	23.37
		Second	-7.85	26.42
	Semantic-Categorical	First	11.07	32.35
		Second	5.33	41.14
Second	Formal-Associative	First	10.11	29.61
		Second	-1.69	20.46
	Semantic-Categorical	First	1.87	44.65
		Second	15.96	19.04

Table 5.3: Mean priming sizes across conditions for pilot group (experiment 2).



Plot 5.2: Mean priming sizes per condition for pilot group. Error bars represent +1 SEM (experiment 2).

5.3.6. Discussion

The pilot study above demonstrated that there were no significant changes in priming size between the two sessions or between the two phases. This was important to have in mind when interpreting the results yielded by the main TMS experiment reported below.

5.4. TMS Experiment (Experiment 3)

The TMS experiment was conducted in the Institute of Cognitive Neuroscience (University College, London), hosted by Dr. Neil Muggleton (May-August 2009).

5.4.1. Stimuli

The same stimuli were used as those in the first and second pilot experiments (sections 5.2.1 and 5.3.1).

5.4.2. Procedure

The same procedure was employed as that in the second pilot experiment (section 5.3.2), but with the involvement of TMS. Participants completed the first two miniblocks, received 40 seconds of cTBS on one of the two CB sites, and, after a 5-minute break, completed the last two miniblocks. Each miniblock lasted approximately 7 minutes. Participants could take a 2-minute break after the completion of each miniblock. The whole session thus lasted 45-50 minutes.

5.4.3. TMS Setup

This TMS study involved stimulation of the right neocerebellar vermis (NCBV), and of a right posterolateral cerebellar (PLCB) site. The vermal lobules of interest here, i.e. VI and VII, have been successfully stimulated in Hashimoto and Ohtsuka (1995), applying the scalp-based coordinates of 1 cm below the inion and 1 cm laterally to the right. The vermis is one of the most accessible structures for CB TMS (Miall & Christensen, 2004), and the superior posterior CB compartments are those closest to the surface of the scalp (see section 4.4.8). These lobules are included in areas 2 and

especially 3 in figure 5.1 below, and, importantly, comprise the NCB compartment of the vermis (see also figure 1.3). Scalp coordinates of 1 cm below the inion and 2 cm laterally to the right have been considered to correspond to a paravermal site, or to a CB hemispheric site close to the paravermis (Miall & Christensen, 2004), and were thus not preferred.

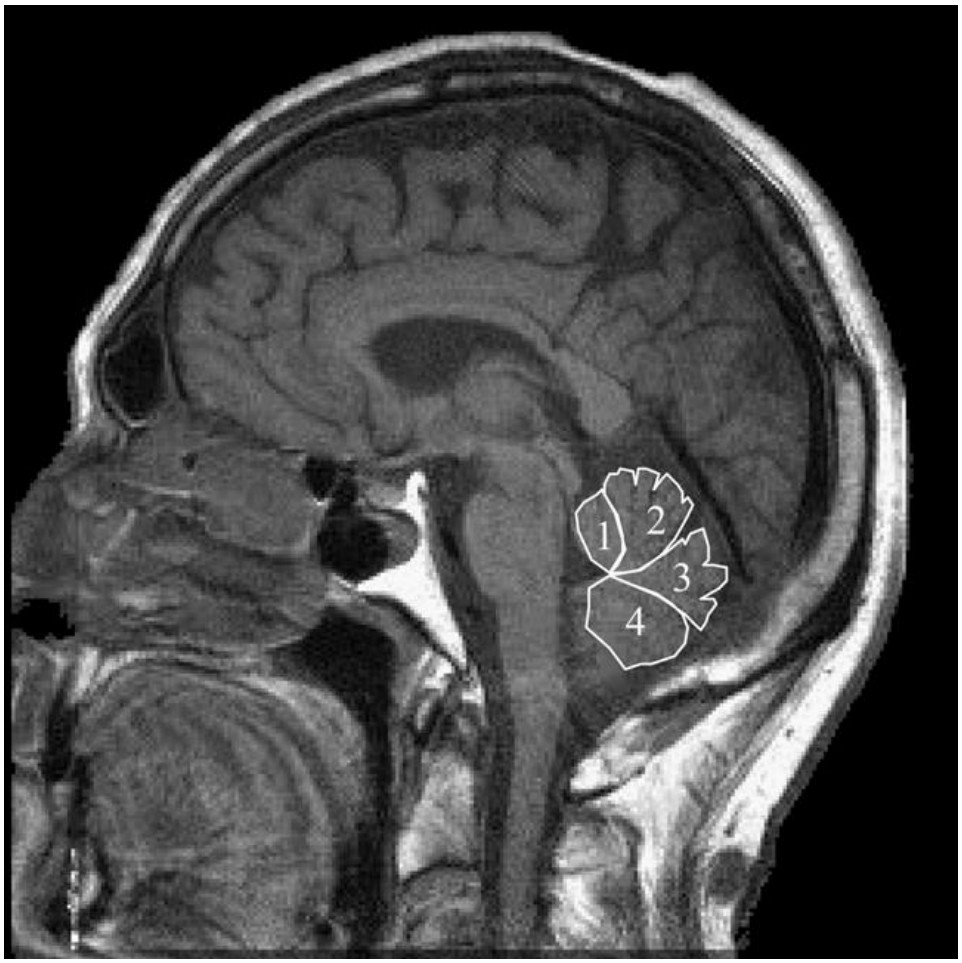


Figure 5.1: Quadripartite division of the vermis by MacLullich *et al.* (2004); area 1: lingula and lobulis centralis (lobules I-III of Larsell & Jansen, 1972); area 2: culmen, bordered by the preculminate fissure and the primary fissure (lobules IV and V); area 3: declive, folium, and tuber, bordered by the primary and the prepyramidal fissure (lobules VI and VII); area 4: pyramis, uvula, and nodulus (lobules VIII-X). Vermal areas 2 and particularly 3 are part of the NCB; they are apparently closer to the scalp and thus more exposed to the TMS coil. Figure reproduced from MacLullich *et al.* (2004, p. 345) with permission © 2004 Elsevier.

With respect to the lateral site of stimulation, language-specific activations are primarily located in HVI (see figure 4.1 in chapter 4). The scalp coordinates of 1 cm below the inion and 3 cm to the right have been amply used in the CB TMS literature (Théoret *et al.*, 2001; Torriero *et al.*, 2004, 2007; Koch *et al.*, 2007, 2008; Miall *et al.*, 2007; Miall & King, 2008). They have been suggested to averagely correspond to parts of right lobules HV and HVI, their cerebral cortical afferents including the right-hand area of the left motor cortex (Koch *et al.*, 2008; Miall & King, 2008). MRI reconstruction and neuronavigation systems in previous studies have confirmed that CB TMS in this site predominantly targets the posterior and superior lobules of the lateral CB (Del Olmo *et al.*, 2007; Koch *et al.*, 2007). Given that the phylogenetically elder, dorsal part of the dentate nucleus projects to motor areas (including primary motor cortex), and that the ventral, phylogenetically recent part projects to prefrontal and premotor areas (e.g. Middleton & Strick, 1998), a more lateral CB site was selected. This site was defined only 1.5 cm further laterally to the right (thus, 1 cm below the inion, and 4.5 cm laterally to the right), since, in terms of straight line distance to scalp, more lateral placement would grossly translate into further distance from the CB cortex (C. Miall, personal communication, 16 April 2009).

In order to estimate the depth of each site and thus the possibility of their successful stimulation, a volunteer was recruited, whose brain image was already registered with the Brainsight™ TMS-MRI co-registration system (Rogue Research, Montreal, Canada).[77] The scalp-based coordinates were transformed into Talairach brain coordinates. For the NCBV site, transformation yielded the coordinates of x = 9 mm, y = -87 mm, z = -21 mm (figure 5.2). The medial site is thus successfully localized below the primary fissure and above the prepyramidal fissure below, and corresponds to the right vermal lobules VI and VII: declive, folium, tuber, i.e. part of the superior posterior cerebellar vermis.

[77] Thanks to Dr. Neil Muggleton for volunteering and helping with the registration and transformation process.

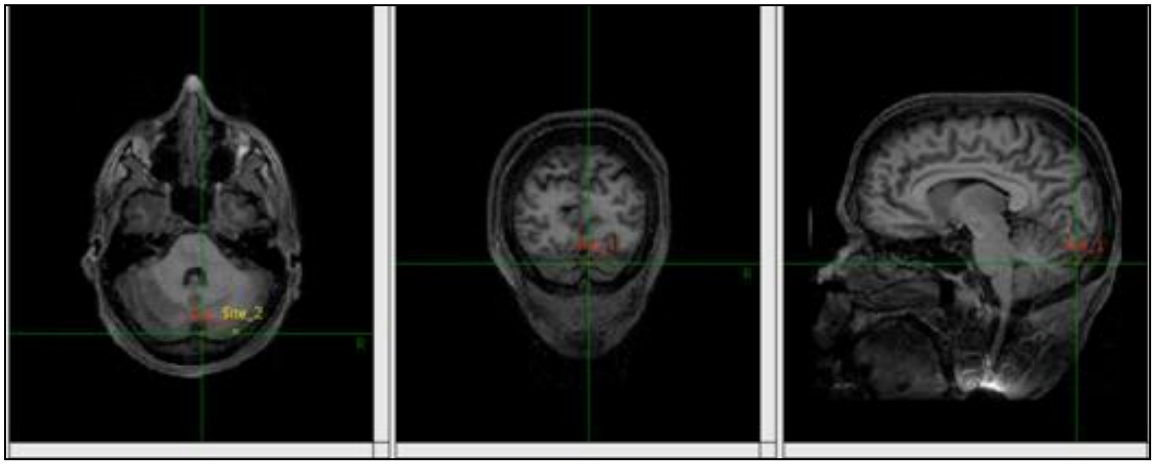


Figure 5.2: Localization of the right NCBV site. Image taken from a Brainsight ™ window, showing the CB area underlying the position of 1 cm below the inion and 1 cm laterally to the right in a volunteer.

For the PLCB site, the transformation yielded the coordinates of x = 33 mm, y = -87 mm, z = -27 mm, corresponding to a significantly deeper right posterior superior hemispheric area, located between the same fissures (figure 5.3).

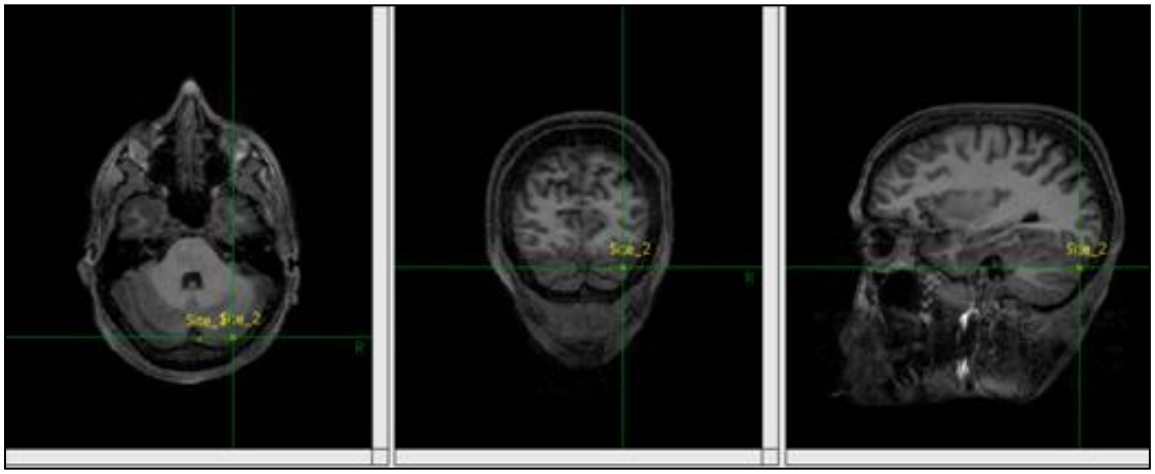


Figure 5.3: Localization of the right PLCB site. Image taken from Brainsight ™ window showing the CB area underlying the position of 1cm below the inion and 4.5 cm laterally to the right in a volunteer.

The cTBS protocol was applied according to the description in section 4.4. Ethical approval was granted by the local ethics committee. All subjects provided their informed consent. All subjects but one (who found the TBS session painful and was not subject to the full 40 seconds, and thus discarded) tolerated TMS well, and all reported a mild discomfort due to muscle twitching.

5.4.4. Participants

All subjects (n = 8) were recruited according to the criteria in section 4.4.12 (mean age: 26.9 years; min: 19; max: 43; SD: 8.6). They all underwent TMS in alternate sites (PLCB, NCBV) in two different sessions, performing the exact same task. A mean time distance of 13.5 days (min 3; max 26 days; SD: 9.1) separated the two sessions. A considerable number of subjects were not recalled for the second session, or were automatically discarded from the analysis, due to certain violations of the protocol: five subjects were not native English speakers despite enrolling as such; in one session, a mechanical error occurred in triggering the TMS machine; in another session, the 5-minute break after the stimulation was not abided by; finally, five subjects failed to reappear for the second session.

5.4.5. Design

For the analysis of TMS effects on priming sizes, a 2 x 2 x 2 within-subjects design was followed: participants were stimulated on two different CB sites: on the right NCBV and on the right PLCB site (site). Priming sizes were calculated before and after TMS (phase). Two different types of priming were examined (priming type): formal-associative (e.g. ‘gift-HORSE’), and semantic-categorical (e.g. ‘table-FURNITURE’).

For the analysis of TMS effects on sensorimotor performance, a 2 x 2 x 2 within-subjects design was also followed: Latencies and accuracy rates were analyzed for both words and nonwords ('wordness') for the phase before and after TMS (phase) of either the PLCB or NCBV (site).

For the TMS effects on processing repeated trials, a similar 2 x 2 x 2 design was followed. The difference in the speed and accuracy of lexical decisions between the first and second (repeated) miniblock of the block before TMS was compared with the same difference after TMS (phase) of the NCBV or of the PLCB (site), both for words and nonwords (wordness).

Full counterbalancing was achieved for the order in which the two sites were stimulated, and for the two stimulus lists of the design (section 5.2.1).[78]

5.4.6. Results

5.4.6.1. Effects on Priming Size

5.4.6.1.1. Lexical Decision Latencies

Trials were excluded according to the criteria in section 4.4.1, and amounted to 1.2% of the data. Mean priming sizes per condition are shown below in table 5.4 and plot 5.3. The analysis demonstrated a selective enhancement of formal-associative priming after NCBV TMS.

[78] As mentioned above, the design of the experiment did not allow for effects on priming sizes to be calculated by items An items-based analysis was only possible by using RTs as a dependent measure. This analysis is presented at the end of appendix A.

A **3-way ANOVA (priming type, site, phase)** only showed **a strong site*priming type*phase interaction (*F* (1, 6) = 20.25, MSe = 331.10, *p* < .005**; also, a non-significant site*phase interaction: *F* (1, 6) = 2.25, MSe = 191.10, *p* = .18; rest of *p*s, *p* > .2). This was primarily owed to the selective increase in formal-associative priming size after NCBV TMS.

2-way ANOVAs: Formal-Associative Priming (site, phase): An analysis for formal-associative priming sizes showed no difference between sites or phases (both *F*s, *F* < 1), but only **a site*phase interaction (*F* (1, 6) = 19.10, MSe = 275.72, *p* < .01)**. As observed in plot 5.3 below, this was because of the selective increase in formal-associative priming size after NCBV TMS,

2-way ANOVAs: Semantic-Categorical Priming (site, phase): An analysis for semantic-categorical priming sizes also demonstrated a site*phase interaction (*F* (1, 6) = 7.58, MSe = 246.47, *p* < .05; rest of *F*s, *F* < 1), which relied, though, on marginal effects (see below).

2-way ANOVAs: before TMS (site, priming type): Priming sizes before TMS did not differ across the two sites (site: *F* < 1). A marginal site*priming type interaction (*F* (1, 6) = 4.34, MSe = 697.81, *p* = .08) was noted though, which once again relied on only marginal differences (see below).

2-way ANOVAs: after TMS (site, priming type): For priming after stimulation, sizes after NCBV TMS were overall larger than those after PLCB TMS (site: *F* (1, 6) = 8.66, MSe = 118.73, *p* < .05), but even more so for formal-associative priming **(site*priming type: *F* (1, 6) = 8.13, MSe = 454.36, *p* < .05)**.

2-way ANOVAs: NCBV conditions (priming type, phase): An analysis of priming sizes in NCBV conditions showed a selective increase for formal-associative

priming **(priming type*phase interaction: *F* (1, 6) = 14.31, MSe = 381.91, *p* < .01)**, and not independently of priming type (phase: $F < 1$).

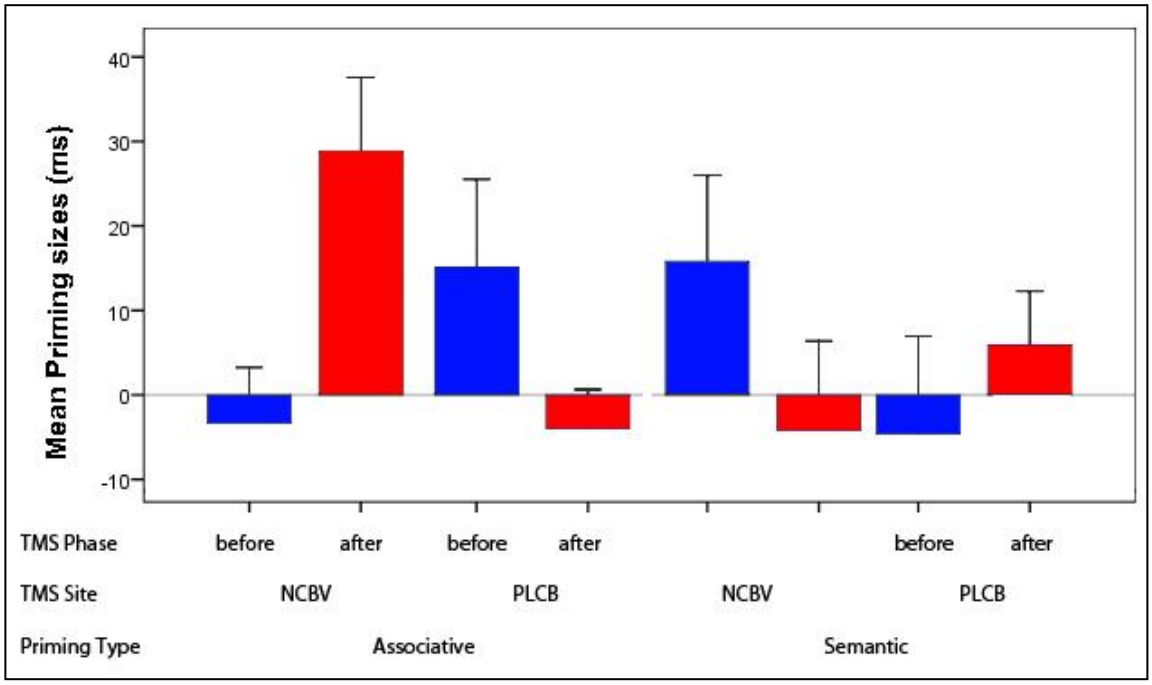
2-way ANOVAs: PLCB conditions (priming type, phase): An analysis of priming sizes in PLCB conditions showed no change independently of priming type (phase: $F < 1$), and an only marginal priming type*phase interaction (F (1, 6) = 4.62, MSe = 379.76, p = .08), which did not rely, though, on any significant effects (see below).

1-way ANOVAs (phase): An analysis of **formal-associative priming** sizes in **NCBV** TMS conditions showed **a significant increase after TMS (phase: *F* (1, 6) = 7.64, MSe = 543.73, *p* < .05)**. No significant change was introduced for **formal-associative priming** after **PLCB** TMS (phase: F (1, 6) = 2.41, MSe = 604.98, p = .17). For **semantic-categorical priming** sizes, there was a marginal decrease after **NCBV** TMS (phase: F (1, 6) = 5.02, MSe = 320.46, p = .07), and no change after **PLCB** TMS (phase: $p > .3$).

1-way ANOVAs (site): Formal-associative priming size after NCBV TMS was also **much larger than that after PLCB TMS (site: *F* (1, 6) = 21.19, MSe = 203.38, *p* < .005)**. Sizes did not differ significantly between the two sites for **formal-associative priming before TMS** (site: F (1, 6) = 2.73, MSe = 583.89, p = .14) or for semantic-**categorical priming before** (F (1, 6) = 3.01, MSe = 554.44, p = .13), or **after TMS** ($p > .3$).

Priming Type	TMS Site	TMS Phase	Mean priming sizes (ms)	SD (ms)
Formal-Associative	**NCBV**	**Before TMS**	-3.34	18.62
		After TMS	28.88	24.70
	PLCB	**Before TMS**	15.15	29.24
		After TMS	-3.95	12.98
Semantic-Categorical	**NCBV**	**Before TMS**	15.84	28.70
		After TMS	-4.21	30.04
	PLCB	**Before TMS**	-4.58	32.52
		After TMS	5.94	17.89

Table 5.4: Mean priming sizes per condition for TMS sessions (experiment 3).



Plot 5.3: Mean priming sizes per condition for TMS sessions. Error bars represent + 1 SEM (experiment 3).

5.4.6.1.2. Lexical Decision Accuracy

Priming sizes were calculated here as the difference between mean accuracy rates for unrelated items and those for related items per condition (4.4.11). The analysis showed a selective decrease in size for formal-associative priming after TMS in the first session of participation (yet independently of site). However, subsequent comparisons with no-TMS groups demonstrated that the change did not reflect a TMS effect, but a general trend in task performance (see Appendix A for the full analysis).

5.4.6.2. Effects on Sensorimotor Performance

5.4.6.2.1. Lexical Decision Latencies

The mean latencies per condition for words and nonwords before and after stimulation of the two sites are shown below in table 5.5 and in plot 5.4. The analysis showed a tendency for RTs to show a smaller decrease after NCBV TMS than after PLCB TMS, selectively for word targets.

A **3-way ANOVA (site, phase, wordness)** showed a clear reduction of RTs in the second phase of each session (phase: F_1 (1, 6) = 11.57, MSe = 1252.43, $p < .05$; F_2 (1, 296) = 124.66, MSe = 2077.07, $p < .001$). Words received significantly shorter latencies across sessions (wordness: F_1 (1, 6) = 8.27, MSe = 4749.98, $p < .05$; F_2 (1, 296) = 71.09, MSe = 7240.31, $p < .001$). Crucially, the analysis also showed a significant by-items **site*phase*wordness interaction (F_1 (1, 6) = 2.93, $p = .14$; F_2 (1, 296) = 6.27, MSe = 1890.44, $p < .05$)**, due to a tendency for RTs to decrease poorly after NCBV TMS for word targets (also, a non-significant wordness*phase interaction: $F_1 < 1$; F_2 (1, 296) = 2.47, MSe = 2077.07, $p = .12$; rest of ps, $p > .2$).

2-way ANOVAs: word targets (site, phase): An analysis for word targets showed the expected reduction of RTs in the second phase, marginally by subjects and significantly by items (phase: F_1 (1, 6) = 4.66, MSe = 1079.20, p = .07; F_2 (1, 148) = 45.68, MSe = 2092.28, p < .001; site: both *F*s, F < 1). Crucially, it also showed a significant by-items **site*phase interaction (F_1 (1, 6) = 1.3, MSe = 593.65, p > .3; F_2 (1, 148) = 7.65, MSe = 1623.68, p < .01)**. As seen in the plot below, this was due to a smaller reduction of RTs after NCBV TMS than after PLCB TMS for word targets.

2-way ANOVAs: nonword targets (site, phase): On the contrary, an analysis for nonwords showed no such interaction (site*phase: both *F*s, F < 1), but only the expected reduction of RTs across sites in the second phase (F_1 (1, 6) = 9.64, MSe = 1022.83, p < .05; F_2 (1, 148) = 81.71, MSe = 2061.86, p < .001). Nonword latencies did not differ between the two sites overall (site: F_1 < 1; F_2 (1, 148) = 2.20, MSe = 2146.58, p = .14).

2-way ANOVAs: NCBV conditions (phase, wordness): A comparison of latencies for NCBV TMS conditions showed the expected reduction in RTs in the second phase (F_1 (1, 6) = 13.27, MSe = 461.83, p < .05; F_2 (1, 296) = 62.66, MSe = 1689.23, p < .001), and the expected difference in latencies between words and nonwords (F_1 (1, 6) = 12.05, MSe = 1483.36, p < .05; F_2 (1, 296) = 49.01, MSe = 4576.25, p < .001). Importantly, though, it also showed a significant by-items **wordness*phase interaction (F_1 (1, 6) = 1.6, MSe = 649.17, p > .2; F_2 (1, 296) = 9.64, MSe = 1689.23, p < .005)**, with a smaller reduction of RTs for words than for nonwords in the phase after NCBV TMS.

2-way ANOVAs: PLCB conditions (phase, wordness): No such interaction was observed in PLCB TMS conditions (wordness*phase: both *F*s, F < 1), apart from the marginal by-subjects and significant by-items main effect of wordness (F_1 (1, 6) = 5.49, MSe = 3921.08, p = .06; F_2 (1, 296) = 65.42, MSe = 4474.84, p < .001), and that of

phase, significantly by items and marginally by subjects (F_1 (1, 6) = 4.76, MSe = 1773.98, p = .07; F_2 (1, 296) = 68.24, MSe = 2278.29, p < .001).

2-way ANOVAs: conditions after TMS (site, wordness): Comparing latencies in the conditions after TMS, words once again received shorter latencies than nonwords (wordness: F_1 (1, 6) = 12.97, MSe = 1224.24, p < .05; F_2 (1, 296) = 45.76, MSe = 4557.50, p < .001). Latencies after NCBV TMS were not larger than those after PLCB TMS overall (site: both *F*s, F < 1), but only selectively so for words; this was shown by a significant by-items **site*wordness interaction (F_1 (1, 6) = 1.6, MSe = 332.77, p > .2; F_2 (1, 296) = 6.81, MSe = 1799.73, p < .05)**.

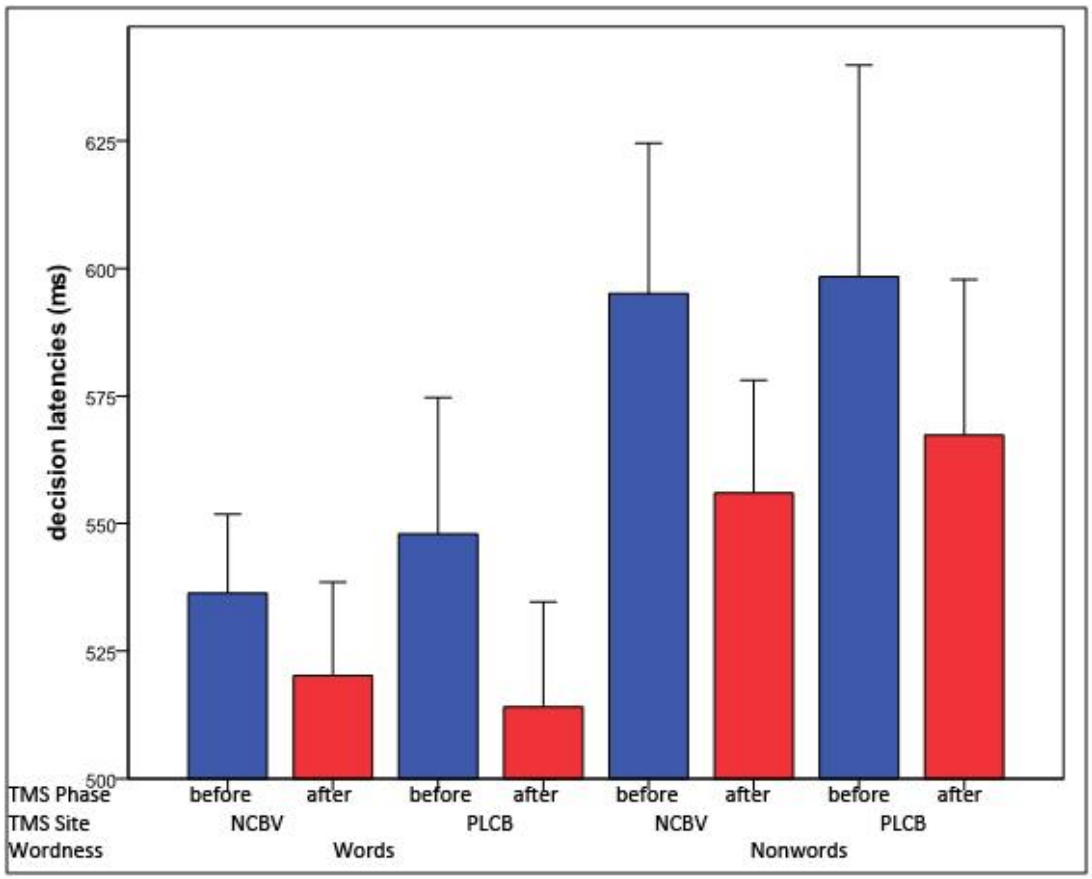
2-way ANOVAs: conditions before TMS (site, wordness): No such interaction was shown for the conditions before TMS (wordness*site: both *Fs*, F < 1), apart from a significant by-items main effect of wordness (F_1 (1, 6) = 5.45, MSe = 4375.34, p = .06; F_2 (1, 296) = 65.40, MSe = 4759.88, p < .001). Latencies between the two sites did not differ before TMS (site: F_1 < 1; F_2 (1, 296) = 2.54, MSe = 1901.49, p = .11).

1-way ANOVAs (for words): phase: Suggestively, the decrease in latencies for words after **PLCB TMS** (phase: F_1 (1, 6) = 4.04, MSe = 1142.68, p = .09; F_2 (1, 148) = 46.45, MSe = 1904.28, p < .001) showed larger effect sizes than for those after **NCBV TMS (phase: F_1 (1, 6) = 1.6, MSe = 530.18, p > .2; F_2 (1, 148) = 10.79, MSe = 1811.68, p < .005)**.

1-way ANOVAs (for words): site: However, the interactions above were at least **partly due to the significantly by-items longer latencies before PLCB than those before NCBV TMS (site: F_1 < 1; F_2 (1, 148) = 4.52, MSe = 1406.68, p < .05). Latencies after NCBV TMS were but marginally longer (by items) than those after PLCB TMS (site: F_1 < 1; F_2 (1, 148) = 3.59, MSe = 1691.97, p = .06).** The analysis was thus inconclusive.

Target Type	TMS Site	TMS Phase	Mean Latencies (ms)	SD (ms)
Words	NCBV	Before TMS	536.33	43.84
		After TMS	520.14	51.94
	PLCB	Before TMS	547.97	75.61
		After TMS	514.02	58.20
Nonwords	NCBV	Before TMS	595.10	83.35
		After TMS	555.93	62.75
	PLCB	Before TMS	598.37	117.25
		After TMS	567.32	86.11

Table 5.5: Mean latencies per condition for TMS sessions (experiment 3).



Plot 5.4: Mean latencies per condition for TMS sessions. Error bars represent + 1 SEM (experiment 3).

5.4.6.2.2. Lexical Decision Accuracy

Accuracy rates per condition were also analyzed, shown below in plot 5.5 and table 5.6. **A three-way analysis** of accuracy rates (wordness, site, phase), without including the 'session number', showed a significant by-items main effect of site, with more mistakes occurring in NCBV conditions overall (site: $F_1 < 1$; F_2 (1, 296) = 4.03, MSe = 0.05, $p < .05$), and a marginal by-items wordness*phase difference ($F_1 < 1$; F_2 (1, 296) = 2.86, MSe = 0.05, $p = .09$; non-significant wordness*site*phase interaction: F_1 (1, 6) = 3.30, MSe = 0.00, $p = .12$; $F_2 < 1$; wordness: F_1 (1, 6) = 2.56, MSe = 0.00, $p = .16$; $F_2 < 1$; rest of *F*s, $F < 1$). However, the analysis for either before or after TMS in particular showed no wordness*site interactions (all *F*s, $F < 1$).

Factoring in the session number in the analysis showed a selective decrease of accuracy rates after NCBV TMS in the first session of participation.

A **4-way ANOVA (site, phase, session number, wordness)** showed a clear learning effect, with significantly less mistakes occurring in the second session of participation (session number: F_1 (1, 4) = 9.72, MSe = 0.01, $p < .05$; F_2 (1, 296) = 5.93, MSe = 0.06, $p < .05$). There also was a significant by-items difference between the two sites, with more mistakes occurring in NCBV conditions ($F_1 < 1$; F_2 (1, 296) = 5.75, MSe = 0.07, $p < .05$). Crucially, however, the analysis showed **a significant by-subjects site*phase*session number interaction (F_1 (1, 4) = 20.28, MSe = 0.00, $p < .05$; F_2 (1, 296) = 1.73, MSe = 0.09, $p = .19$)**. As it will be shown below, this was due to the lower accuracy rates in the condition after NCBV TMS in the first session, independently of target type (site*phase*session number*wordness: both *F*s, $F < 1$). The analysis also showed a number of interactions that are not further discussed, for the sake of brevity (wordness*session number: F_1 (1, 4) = 10.33, MSe = 0.01, $p < .05$; $F_2 < 1$; site*session number: $F_1 < 1$; F_2 (1, 296) = 8.21, MSe = 0.09, $p < .005$; phase*wordness: $F_1 < 1$; F_2 (1, 296) = 4.16, MSe = 0.09, $p < .05$; site*session number*wordness: $F_1 < 1$; F_2 (1, 296) = 6.64, MSe = 0.09, $p < .05$; wordness*site*phase interaction: F_1 (1, 4) = 5.82, MSe =

0.00, p = .07; F_2 (1, 296) = 2.25, MSe = 0.08, p = .14; session number*phase: F_1 (1, 4) = 1.76, MSe = 0.00, p > .25; F_2 (1, 296) = 2.23, MSe = 0.07, p = .14; rest of Fs, F < 1, or p > .2).

The **non-selective nature of this site*phase*session number interaction for words or nonwords** is demonstrated in the two separate ANOVAs for words and nonwords below:

3-way ANOVAs: Words (site, phase, session number): An analysis for accuracy rates for word targets showed **a significant by-subjects site*phase*session number interaction (F_1 (1, 4) = 13.96, MSe = 0.00, p < .05; F_2 < 1).** The analysis also showed a marginal by-items site difference, with higher accuracy rates for NCBV conditions overall (F_1 (1, 4) = 2.04, MSe = 0.00, p > .2; F_2 (1, 148) = 3.07, MSe = 0.08, p = .08), and a significant by-subjects learning effect (session number: F_1 (1, 4) = 26.12, MSe = 0.01, p < .01; F_2 (1, 148) = 2.75, MSe = 0.07, p = .1; also: site*session number: F < 1; F_2 (1, 148) = 17.30, MSe = 0.08, p < .001; phase*session number: F_1 (1, 4) = 3.28, MSe = 0.00, p = .14; F_2 (1, 148) = 1.10, MSe = 0.06, p > .3; rest of Fs, F < 1, or p > .25).

3-way ANOVAs: Nonwords (Site, phase, session): An analysis for nonword targets **showed the same, significant by-subjects site*phase*session number interaction (F_1 (1, 4) = 17.71, MSe = 0.00, p < .05; F_2 (1, 148) = 1.08, MSe = 0.09, p > .3)**, along with a number of marginal or non-significant effects (session number: F_1 < 1; F_2 (1, 148) = 3.21, MSe = 0.06, p = .08; phase: F_1 < 1; F_2 (1, 148) = 3.11, MSe = 0.10, p = .08; site*phase: F_1 (1, 4) = 2.89, MSe = 0.00, p = .16; F_2 (1, 148) = 1.12, MSe = 0.08, p > .25; site: F_1 < 1; F_2 (1, 148) = 2.68, MSe = 0.07, p = .1; rest of Fs, F < 1, or p > .25).

Subsequent analysis including the wordness factor showed that the effects in two- and one-way ANOVAs did not reach significance separately for words or for

nonwords. Thus, for the sake of brevity, and on the grounds of the above non-selective nature of the site*session number*phase interaction with respect to wordness (word/nonword), the following analysis is reported instead, collapsing on the wordness distinction.

2-way ANOVAs: NCBV conditions (phase, session number): An analysis for NCBV TMS conditions showed a significant by-subjects and marginal by-items **phase*session number interaction (F_1 (1, 4) = 11.28, MSe = 0.00, p < .05; F_2 (1, 296) = 3.35, MSe = 0.08, p = .07)**, with more mistakes occurring selectively after NCBV TMS for the first session. No effect of NCBV TMS on accuracy was observed independently of session number (phase: both *F*s, $F < 1$). The analysis also showed an expected session number effect, with fewer mistakes occurring for the second sessions (F_1 (1, 4) = 6.06, MSe = 0.01, $p < .05$; F_2 (1, 296) = 13.45, MSe = 0.09, $p < .005$).

2-way ANOVAs: PLCB conditions (phase, session number): No such interaction was observed for PLCB TMS conditions (both *F*s, $F < 1$); there was only a significant by-subjects learning effect (session number: F_1 (1, 4) = 7.63, MSe = 0.00, $p < .05$; F_2 (1, 296) = 1.2, MSe = 0.09, $p > .3$).

2-way ANOVAs: After TMS (site, session): Comparing the accuracy rates after TMS of the two sites showed **a significant site*session interaction (F_1 (1, 4) = 8.20, MSe = 0.00, p < .05; F_2 (1, 296) = 9.74, MSe = 0.08, p < .005)**, with lower accuracy rates in the condition after NCBV TMS in the first session. Accuracy rates after TMS did not differ between the two sites independently of session number (site: both *p*s, $p > .2$). As above, there was a session number effect, with fewer mistakes occurring in the second session of participation (F_1 (1, 4) = 8.85, MSe = 0.01, $p < .05$; F_2 (1, 296) = 6.02, MSe = 0.06, $p < .05$).

2-way ANOVAs: Before TMS (site, session): On the contrary, comparing accuracy rates before TMS of the two sites, no such interaction was observed

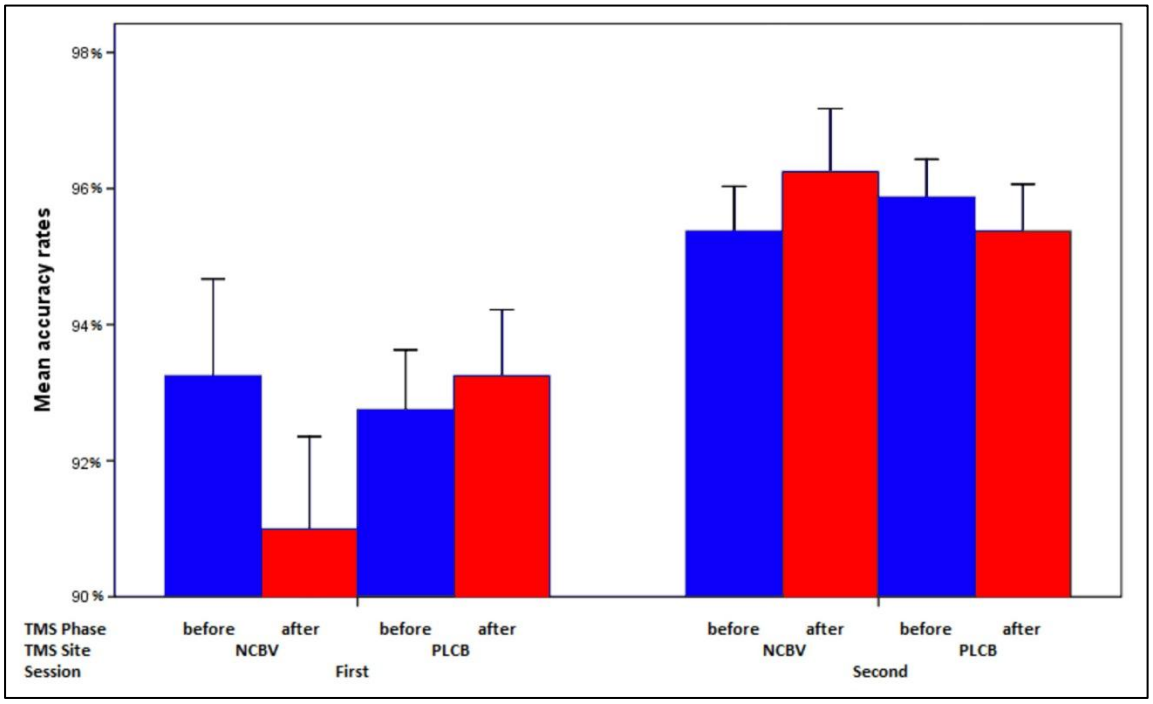
(site*session: both *p*s, $p > .25$). Instead, there was a significant by-items effect of site, with NCBV conditions involving higher accuracy rates before TMS (site: $F_1 < 1$; F_2 (1, 296) = 4.00, MSe = 0.07, $p < .05$). Accuracy rates in the second session were marginally higher by subjects (session: F_1 (1, 4) = 5.03, MSe = 0.00, $p = .07$; $F_2 < 1$).

2-way ANOVAs: first session (site, phase): An analysis for the first session of participation showed a significant by-subjects decrease in accuracy rates after TMS (phase: F_1 (1, 2) = 13.9, MSe = 0.00, $p < .05$; $F_2 < 1$), which was stronger after NCBV TMS – this was demonstrated by a significant by-subjects **site*phase interaction (F_1 (1, 2) = 9.95, MSe = 0.00, $p < .05$; F_2 (1, 296) = 1.2, MSe = 0.06, $p > .25$)**. NCBV conditions also involved lower accuracy rates overall, significantly so by items (F_1 (1, 2) = 1.2, MSe = 0.00, $p > .3$; F_2 (1, 296) = 16.17, MSe = 0.00, $p < .001$).

2-way ANOVAs: second session (site, phase): Unexpectedly, a comparison of accuracy rates of participants in their second sessions showed a significant by-subjects site*phase interaction (F_1 (1, 2) = 25.62, MSe = 0.00, $p < .05$; $F_2 < 1$; phase: $F_1 < 1$; F_2 (1, 296) = 1.9, MSe = 0.06, $p = .17$; site: both *F*s, $F < 1$). However, this interaction did not rely on any significant effects of TMS in the second session, as seen below.

1-way ANOVAs (phase): A comparison of accuracy rates before and after TMS for participants that underwent **NCBV TMS for the first session** showed **a strong effect of phase by subjects (F_1 (1, 2) = 62.50, MSe = 0.00, $p < .005$; F_2 (1, 296) = 1.3, MSe = 0.08, $p > .3$)**, with accuracy rates dropping after NCBV TMS for the first session. On the contrary, no significant change was introduced in accuracy rates after **NCBV TMS in the second session** (phase: F_1 (1, 2) = 1.6, MSe = 0.00, $p > .25$; F_2 (1, 296) = 2.80, MSe = 0.05, $p = .1$). No change was introduced in accuracy rates after PLCB TMS in the **first** or in the **second session** (all *F*s, $F < 1$).

1-way ANOVAs (site): Similarly, the accuracy rates in the condition **after NCBV TMS** in the **first session** were significantly lower by items than those **after PLCB TMS for the first (F_1 (1, 4) = 1.8, p > .2; F_2 (1, 296) = 9.21, MSe = 0.10, p < .005) or for the second session (F_1 (1, 4) = 9.36, MSe = 0.00, p < .05; F_2 (1, 296) = 5.07, MSe = 0.08, p < .05).**



Plot 5.5: Mean accuracy rates per condition for TMS sessions. Error bars represent +1 SEM (experiment 3).

Site	Phase	Session	Mean accuracy rates (%)	SD (%)
NCBV	**Before**	**1**	93.25	2.84
		2	95.38	1.32
	After	**1**	91.00	2.71
		2	96.25	1.85
PLCB	**Before**	**1**	92.75	1.76
		2	95.88	1.11
	After	**1**	93.25	1.94
		2	95.38	1.38

Table 5.6: Mean accuracy rates per condition for TMS sessions (experiment 3).

5.4.6.3. Effects on Practice-Induced Facilitation

No changes were observed in the speed or accuracy of lexical decisions in repeated trials selectively after TMS of a particular site (see section 5.4.5 for design).

5.4.6.3.1. Lexical Decision Latencies

A three-way ANOVA on differences in latencies showed significantly larger by-items facilitation sizes for NCBV conditions across phases (site: F_1 (1, 4) = 3.64, MSe = 2756.12, p = .1; F_2 (1, 296) = 8.59, MSe = 3353.15, p < .005), along with a decrease in facilitation after TMS across sites, significantly by items (phase: F_1 (1, 4)= 1.19, MSe = 3555.12, p > .3; F_2 (1, 296) = 6.93, MSe = 3312.03, p < .01; rest of *F*s, F < 1). However, a comparison of the TMS group with the pilot group above (section 5.3), collapsing on the site variable, yielded no differences between the two (all *F*s, F < 1). This demonstrated that the changes after TMS were not due to a TMS effect.

5.4.6.3.2. Lexical Decision Accuracy

Similarly, no changes in practice-related effects could be shown after TMS of any particular site. The same three-way ANOVA showed a wordness*site*phase interaction by subjects (F_1 (1, 4) = 11.56, MSe = 0.39, p < .05; F_2 < 1), and a marginal by-items site*phase interaction (F_1 (1, 4) = 2.2, MSe = 0.39, p > .2; F_2 (1, 296) = 2.96, MSe = 0.11, p = .09). However, the interaction was due to differences between the two sites before TMS (site: F_1 (1, 4) = 15.36, MSe = 0.69, p < .05; F_2 (1, 296) = 4.07, MSe = 0.11, p < .05), and not to the phase after TMS (site, site*wordness: all *p*s, p > .2). Including the session number in the analysis yielded effects of the same size and significance. Collapsing on the site variable, no differences were found between the TMS group and the pilot group (all *F*s, F < 1).

5.4.7. Discussion

The results above demonstrated a selective drop in accuracy rates after NCBV TMS in the first session of participation, and also a tendency for longer latencies for word responses after NCBV TMS (section 5.4.7.1). Above all, they showed a selective enhancement of formal-associative priming after NCBV TMS, while TMS on the PLCB site provided adequate control conditions. No priming-related changes in accuracy rates were captured, and no effects of TMS in processing repeated trials (section 5.4.7.2).

5.4.7.1. Sensorimotor Effects

The results above provided some evidence for a disruption of lexical decision performance after NCBV TMS that could be due to TMS effects on reading and/or button presses.

To begin with, there was a selective drop in accuracy rates after NCBV TMS for the first session of participation. Familiarity (learning effects) with the stimulus set conceivably annulled any such effects when participants underwent NCBV TMS for the second session. However, these effects were not selective for words or nonwords, and could only be demonstrated by collapsing on the wordness (word/nonword) distinction; the number of participants conceivably did not suffice to show significant effects separately for words and nonwords, or to demonstrate whether the effect was stronger for one of the two types.

Vermal pathology is indeed associated with reading deficits. Reading errors have been observed in a group of selectively vermal patients that were expert readers before the onset of the lesion (Moretti *et al.*, 2002). The vermis may support reading in a number of different ways. Briefly, here, visual information conveyed by mossy and climbing fibers is integrated in the vermal Purkinje cells; the flocculonodular lobe and

the vermis modulate the nictitating membrane response (section 2.3.4.3); vermal lesions may cause saccadic dysmetria, incapacitating ocular vergence movements (see Moretti *et al.*, 2002, for more details). Moreover, the posterior vermis in particular accelerates saccades in the ipsilateral horizontal direction (Ohtsuka & Enoki, 1998, for discussion), and TMS in that region briefly disrupts saccades, smooth pursuit eye movements and eye-head gaze shifts (Hashimoto & Ohtsuka, 1995; Ohtsuka & Enoki, 1998; Nagel & Zangemeister, 2003). Thus, given that portions of the right posterior vermis were stimulated here, and that reading in English is performed from left to right, reading may have been disrupted. Another explanation of these mistakes would be that CB TMS selectively disrupted the elementary associations made between buttons and lexical decisions, i.e. 'left-hand button: nonword', 'right-hand button: word'. Evidence was discussed earlier for the significance of the CB in forming elementary associations (section 2.3.4.3). The current design does not allow for a dissociation between the two mechanisms.

Moreover, latencies for word targets did not tend to be as short after NCBV TMS as after PLCB TMS. If such tendencies were observed independently of string type (word/nonword), then the most suitable interpretation would have been that the disruption pertains to visual, rather than to motor processing. The motoric explanation of the disruption is based on the reciprocal connections of the right CB with areas of the contralateral (left) motor and premotor cortex, which are responsible in turn for the control of contralateral (right) limb movement. Thus, the motoric explanation accommodates disruptions selectively for the 'word' response, which is confounded here with right-hand index finger button presses. The tendencies for relatively slow right-hand index finger responses (word strings) after NCBV TMS may in fact require a motor-based explanation. The data do not allow for a dissociation between the two main explanations, and might as well be explained by a combination of limb-motor and saccadic-reading disruptions. More participants would be required to clarify the nature of this trend.

5.4.7.2. Cognitive Effects

5.4.7.2.1. The Neocerebellar Vermis and Associative Priming

The major finding of this study is that NCBV TMS selectively enhanced formal-associative priming. This cannot be explained on the basis of disruptions in reading or in button presses. All such explanations would predict that both targets and controls of both priming types should receive lower accuracy and/ or longer latencies in the post-NCBV TMS conditions. Similarly, there is no reason to believe that this enhancement involved a strategic compensation for some other disruption: the short SOAs and the low relatedness proportion employed make strategic involvement highly unlikely; such compensation would also not explain the lack of concomitant changes for the semantic-categorically related pairs.[79]

This calls us to examine NCBV involvement in cognitive tasks independently of reading mechanisms. The neurofunctional explanation here would be similar to that provided for CB TMS-induced facilitations in other studies. Purkinje cells, which are responsible for inhibiting and dampening the constant excitation that deep CB nuclei undergo (section 1.2.3.1), are highly accessible to TMS. The latter introduces neural noise to the inhibitory activity of this layer in the NCBV, and allows inappropriate levels of excitatory output to the cortical targets (through the thalamus) of the deep CB nuclei (section 4.4.9). The findings here were thus in line with the predictions made in the second chapter, where it was hypothesized that disruption of CB cortical activity would amount to disinhibited predictions in phonological-syntactic processing (section 2.6.2.1). However, the particular neuroanatomical details of CB cTBS are still relatively unexplored to commit to a particular explanation of the direction of the effect here (see section 4.4.9 for discussion). Thus, the discussion below is also strongly motivated by the selective nature of the effect for a priming type (formal-associative priming).

[79] The same effect was shown in a ratio-based analysis (section 4.4.11); thus the change was not because of larger overall RTs after NCBV TMS – this would still not explain the selective nature of the effect for priming type.

The explanation put forth here stands at the intersection of two largely independent strands of research in schizophrenic cognitive-linguistic deficits: one on selective NCBV impairment, and one on abnormally strong lexical priming effects. Correlations between the two have not been extensively studied, but have been reported before. For instance, increased vermal white-matter volume in schizophrenic males has been associated with poorer immediate verbal memory and severity of positive symptoms (Levitt *et al.*, 1999).

Vermal atrophy is the most widely cited structural CB abnormality in schizophrenia (Picard *et al.*, 2008). Brain-imaging research has lately established a selective NCBV volume reduction in schizophrenics (e.g. Martin & Albers, 1995; Ichimiya *et al.*, 2001; Okugawa *et al.*, 2007); interestingly, this is not accompanied by PLCB volume reduction (Okugawa *et al.*, 2003). However, much unlike the PLCB loci, the anatomical details supporting NCBV cognitive functions are poorly investigated. For instance, NCBV-related cognitive deficits in schizophrenics have been recently discussed (Ichimiya *et al.*, 2001) with reference to studies on the connectivity of the vermis with the thalamus, limbic systems, and frontal lobe regions via the fastigial nucleus in the cat (Harper & Heath, 1973). According to other considerations, impaired vermal Purkinje cell inhibitory output to the fastigial nuclei induces overactivity of dopaminergic basal ganglionic neurons, which may in turn deregulate frontal cortical function via the cortico-striatal loop (Martin & Albers, 1995). Indeed, enhanced priming of various types has been independently reported in schizophrenic populations, in both lexical decision and word pronunciation paradigms, and is commonly attributed to a disinhibited activation spread in the semantic network (see Moritz *et al.*, 2001, for a brief review).

However, the effects induced may arguably reflect a direct disruption of NCB-based language functions. There is imaging and clinical evidence implicating the NCBV in non-motor-related linguistic processes (e.g. Desmond *et al.* 1998; section 2.4 here). In that study, for instance, the ‘FEW’ condition of the stem completion task was thought to

reflect NCB (both NCBV and PLCB) constrained search mechanisms for unifications at the sublexical level (e.g. stem: 'PSA'; completion: 'LM'; section 2.4.4.2). By analogy, here, the selective effect of NCBV TMS on formal-associative priming would reflect constrained search mechanisms at the supralexical, phrasal level, between words that co-occur in idiomatic phrases (e.g. prime: 'gift'; target: 'horse'). Most importantly, the selective nature of the TMS effects on formal-associative priming would prompt explanations on the basis of CB associative computations (sections 2.3.4): upon perception of the prime-word, the strong transitional probabilities to the following target word would establish low Kalman gain conditions, inviting maximized involvement of NCB internal models in processing the upcoming target word. On the contrary, if the disruption induced was a product of disinhibited cerebral functions (indirectly caused by CB cortical disruption), then enhancements would have occurred for both formal-associative and semantic-categorical priming.

Thus the study here pertains to selective disruptions of particular lexical priming types – this is very poorly explored in schizophrenia research. The latter is predominantly conducted within the study of thought-disordered types, and predominantly revolves around semantic-categorical priming (e.g. Picard *et al.*, 2008, for a meta-analysis), with poor emphasis on the linguistic aspects of the disruption. On the other hand, the boosted formal-associative priming here reflects a linguistic deficit in a clearer fashion, in the sense of computing the co-occurrence relations of word pairs forming idiomatic phrases. There is indeed some first supportive evidence for a selective impairment in associative priming in schizophrenics. A poorly cited study[80] (Nestor *et al.*, 2006) showed that, while healthy matched controls were primed more for lexical decisions on target words related with primes via two relations (''semantic-and-associated'', e.g. 'doctor-nurse') than for pairs related only semantically (e.g. 'deer-pony') or only associatively (e.g. 'bee-honey'), schizophrenics showed greater priming for associated-only words than for words related only semantically or both semantically

[80] Cited by only 5 articles, according to a 'Google Scholar' search (26 May 2010).

and associatively.[81] This was interpreted on the basis of ‘‘an associational bias that leads to restricted semantic integration and contributes to disturbed thinking’’ (*ibid*, p. 142). In this respect, then, it would be interesting to examine whether NCBV pathology involves ‘associational biases’ as direct effects, and semantic-categorical deficits as only by-products. The latter could be either due to interference from these associational biases, or due to the deregulated cortical operations that disinhibited fastigial output would induce, as discussed above. TMS could offer further insights into such dissociations, by contrastive stimulation of different cerebral and CB loci.

Such disinhibited associative priming would arguably underlie the symptomatology of ‘‘positive disordered discourse’’. This traditionally involves excessive and incoherent speech, and, crucially, tangentiality (i.e. responses related to more superficial aspects of the conversation and not to the essence of the topic) and derailment (i.e. a tendency for discourse to move off track; Andreasen, 1986). In this respect, the ‘schizophrenic-like’ associative boost observed here is reminiscent of the extravagantly tangential, hyper-associative discourse of the spokesman in Monty Python’s ‘‘Word Association Football’’ (Cleese, 1973) cited in the epigram at the start of this chapter.[82] In fact, these discourse impairments have been accounted for in terms of deficient suppression of memory nodes irrelevant to the current discourse (Gernsbacher *et al.*, 1999). A deficient NCBV cortex might thus provide only weak inhibitory signals to the deep CB nuclei, thus allowing for superfluous suboptimal associations to continue activating cerebral representations that are irrelevant with the information currently processed. Thus, the hypothesis developed here about the psycholinguistic/cognitive functions of the NCBV could bring together discourse

[81] As in most studies, Nestor *et al.* (2006) do not clarify whether the associatively related pairs were particularly associated in the formal or semantic level.

[82] The term ‘word association’ refers to the psychological technique invented by Jung in which patients are asked for an immediate response to words given to them. By extension, it also refers to the word game in which the first player starts with a word and subsequent players have to follow with a related word and so on. The term ‘word association football’, on the other hand, involves a disinhibited merger of two different word compounds, that of ‘word association’ and that of ‘association football’.

processing approaches to schizophrenic discourse and neurocognitive considerations on the NCBV.[83]

These findings would seem to support the idea that NCBV Purkinje cells are involved in the inhibition of the expression of particular associations in inappropriate contexts. However, inhibitory Purkinje cell function supports much more than simple behavioral inhibition. Acquisition of associative properties of neural events by the CB is achieved with climbing fiber signals depressing parallel fiber-to-Purkinje cell signalling, thus releasing deep nuclei from Purkinje cell inhibition and allowing output signals from the CB microcomplex, which may be further adapted (sections 1.2.3; 2.3.3). Modification of Purkinje cell activity underlies expression, adaptation, and inhibition of the learned response. Suggestively, the metaphor of ‘‘sculpting’’ (Eccles *et al.*, 1967) has been used to describe the function of the inhibitory output of the CB cortex to the nuclear neurons:

> [l]ike a block of marble, activity within the nuclei is relatively amorphous and theoretically multi-dimensional until it is given form by the inhibitory ‘sculpting’ of the cortical input (Thach, 2007, p. 164).

The findings also have far-reaching consequences for the neural foundations of psycholinguistic mechanisms of sentence processing, as studied in experimental psycholinguistic (chapter 2), or historical linguistic contexts (chapter 3). The involvement of the right NCBV in formal-associative priming promotes its significance in phenomena like articulatory reduction, chunking, and phoneme restoration (see sections 2.6 and 3.6 for references). NCBV predictions would allow for the acceleration of articulatory gestures in producing a sentence with contextually predictable items, and

[83] The involvement of NCBV circuitry in formal-associative priming need not suggest that NCBV circuitry is restricted to a formal level of analysis. Arguably, the lack of coherence in schizophrenic discourse would involve disinhibited semantic associations as well (sections 2.4.4.5 and 2.4.4.6), which are, however, beyond the scope of this TMS study.

would allow for the rapid co-articulation of co-occurring items; the macroscopic effects of these may be seen in phonetic attrition and chunking in the contexts of grammaticalization and idiomatization (section 3.6.3.1). In sentence comprehension and perception, NCBV predictions may be used to restore noisy percepts, on the basis of transition probabilities in a sublexical phonological-syllabic level (Samuel, 1981), or in a supralexical one, as in sentential contexts (Warren, 1970; section 2.6.3.1 here).

5.4.7.2.2. On Covert Speech Production

Conceivably, an alternative interpretation of this enhanced associative priming would be based on CB involvement in ‘‘silent/covert speech’’ (e.g. Ackermann *et al.*, 1998), and on the emulative employment of such processes in language perception-comprehension (Pickering & Garrod, 2007; section 2.3.2 here). While stimulation of right paravermal compartments are more likely to be activated by placing the coil 2 cm laterally to the right from the inion instead of 1cm (e.g. Miall & Christensen, 2004), the scalp coordinates used here for targeting the right NCBV may concomitantly stimulate portions of the right superior posterior paravermis (e.g. Hashimoto & Ohtsuka, 1995). In this case, the premotor cortical outputs of the paravermal afferents via the interpositus nucleus (e.g. Middleton & Strick, 2001) might have affected the language production system, especially its lower phonetic/phonological representation. Characteristic examples would be the ataxic dysarthria observed in paravermal lesions (see Ito, 2000), and the cases of mutism reported after vermal malformations (e.g. Al-Anazi *et al.*, 2001). Moreover, the premotor cortex (one of the output loci of these CB compartments) has been argued to participate in speech perception via the emulative employment of the speech production system (Iacoboni, 2008). Thus, upon perception of the prime word, the strong transitional prime-to-target-word probabilities would establish low Kalman gain conditions, inviting maximized, predictive involvement of the production system in the perception of the target word. In that way, then, disruption of vermal-paravermal function would result in abnormalities in predictive functions, due to impairments in the language production circuit (Pickering & Garrod, 2007).

While this explanation cannot be directly falsified here, assuming that NCB compartments are involved in language perception via their participation in language production stands as both a rather parsimonious and a rather over-specifying assumption. This interpretation would overlook two important aspects of CB function: its domain-general, modality-independent (perception, production) contributions, supported by its massive reciprocal connectivity with all major areas of the CNS; and its widely accepted role in directly grounding feedforward control, state estimation, and/or associative computations (see section 2.5 for further discussion).[84]

5.4.7.2.3. On the Absence of Lateral Cerebellar TMS Effects

The absence of effect after TMS of the PLCB site is interpretable in different ways. To begin with, while the scalp coordinates for the NCBV are precise enough, there has been no strictly language-related TMS study of the PLCB (section 4.4.8). The CB hemispheres constitute a much larger area for stimulation, thus maximizing the range of candidate stimulation coordinates. The PLCB coordinates used here also corresponded to a significantly deeper site: while the medial site had a Talairach z value of -21 mm, the lateral one had a z value of -27 mm. Thus, the possibility that the PLCB is involved in formal-associative relations in the same way as the NCBV can not be excluded. It may thus be that stimulation with more penetrative coils might induce the same enhancing effects observed (see section 8.2.1 for discussion). The same pattern (present medial CB but absent lateral CB TMS effects) has been found before in a finger tapping task, and has received the same interpretation (Théoret *et al.*, 2001). Alternatively, the PLCB might perform the same computations with the NCBV, but with

[84] While further research would be required, the absence of associative priming deficits in Broca's or Wernicke's pathology (Hagoort, 1997) would also suggest that deficits in language production-related loci need not necessarily yield disruptions in language prediction-emulation.

different inputs/outputs, by being involved in semantic associations (section 2.4), but not in formal ones.[85]

5.4.7.2.4. On the Lack of Change of Practice Effects

No effects were found on performance in repeated trials after TMS of any particular site. This would probably suggest that the particular loci stimulated were not relevant with practice-induced facilitation processes. Alternatively, while such impairments are associated with PLCB lesions (e.g. Fiez *et al.*, 1992; Gebhart *et al.*, 2002; section 2.4.2 here), the PLCB site stimulated here might be relevant but not reachable with a conventional focal coil.

[85] As discussed above, the semantic-associative strength that some of the formally associated pairs could have was not directly measured (section 5.2.1). Instead, as all pairs formed idiomatic phrases, they were assessed as formal associates.

5.5. Conclusions

The CB TMS experiment presented above is the first linguistic CB TMS study in the literature; it promotes the association of particular schizophrenic brain pathology with purely linguistic impairments, and, above all, it provides suggestive evidence for the selective involvement of a NCB locus in associative linguistic mechanisms. This is in line with the theoretical framework and the model developed in chapter 2, whereby the neocerebello-cortical circuit performs associative linguistic computations that are adaptively combined with cortico-cortical, categorical ones.

Chapter 6

TMS Study 2: Effects of Theta-Burst Stimulation of the Right Neocerebellar Vermis on Semantic-Categorical Priming

> ‘‘Doctor, I have pains in my chest and hope and wonder if my box is broken and heart is beaten for my soul and salvation and heaven, Amen’’ (schizophrenic patient; in Maher, 1968, November; cited in Forrest, 1976, p. 289)

6.1. Introduction

The first study (chapter 5) showed a selective enhancement of formal-associative priming after TMS of the right neocerebellar vermis (NCBV), whereas the posterolateral cerebellar (PLCB) site provided adequate control conditions. The present study, also reported in Argyropoulos *et al.* (2010), attempted to replicate these findings, using stimuli and native speakers of another language, Modern Greek.

In the first study, half of the stimuli were repeated, in an attempt to capture any differences that cerebellar (CB) TMS might have on the practice-related decrease in RTs and error rates in processing repeated items. However, no such differences were found. Thus, no repeated items were included in the present experiment. This allowed for a greater number of related pairs to be used. It was also possible to recruit a larger subject pool, given the availability of a TMS apparatus in a hospital in the researcher’s home city (‘Papanikolaou’ Hospital, Thessaloniki, Greece). This allowed for a counterbalanced design for both subjects and items (F_1, F_2).

6.2. TMS Experiment (Experiment 5)

The TMS experiment was conducted in the premises of the 3rd Neurological Clinic of the 'Papanikolaou' Hospital in Thessaloniki, Greece, hosted by Prof. V. Kimiskidis and Dr. S. Papagiannopoulos (November-February 2009). Ethical approval was granted by the local ethics committee, and clinical staff was available at all times.

6.2.1. Stimuli

An apparent disadvantage in the construction of the stimulus set was the very unavailability of published free word-generation norms or similarity measures, or of any known Modern Greek lexical decision studies with reported stimulus sets. Whereas these circumstances would normally prompt the experimenter to perform a free word-generation normative study, as in other lexical decision experiments (e.g. Rogers & Friedman, 2008), such a study was not preferred here. This was because immediately co-occurring nouns in the form of [$N_{[nom]}$ $N_{[nom]}$] sequences are already rare in Modern Greek, to the extent that a sizeable set would hardly be obtained from any such word-generation study.

In fact, many idiomatic [N_1N_2] constructions in English are translated in Modern Greek with the amply used [$N_{2[nom]}N_{1[gen]}$] sequence, as in the examples (1a-c) below.

(1a) βόμβα υδρογόνου
'vomva iδro'γonu
bomb.SG.NOM. hydrogen.SG.GEN.
'hydrogen bomb'

(1b) κοιλίδα πετρελαίου
ci'liδa petre'leu
slick.SG.NOM. oil.SG.GEN.
'oil slick'

(1c) φούρνος μικροκυμάτων
'furnos mikroci'maton
oven.SG.NOM microwave.PL.GEN.
‘ microwave oven’

Despite that, $N_{[nom]}$-$N_{[nom]}$ sequences were preferred over $N_{[nom]}$-$N_{[gen]}$ ones, as nominative case was assumed to have a lower recognition threshold than other cases; prime and target words were also attempted to stay equal in all features. Thus, the formal-associatively related items were taken from idiomatic phrases in Modern Greek abiding by the case restriction above. The co-occurring word pairs were allowed to have a range of 0-2 words mediating between them, given the very low frequency of immediately co-occurring N-N sequences of the same case in Modern Greek. On the contrary, the set of semantic-categorically related pairs was constructed on the basis of their lower co-occurrence frequency, and with the same length and word frequency as in the associates set. The stimuli are shown in tables B.1 and B.2 in Appendix B.

The co-occurrence frequency for the words in each pair was determined by the web-pages retrieved by Google.[86] Formal associate pairs were found in the particular idiomatic expression where they co-occurred in an average of 58,340.00 Google pages (min: 3,040.00, max: 316,000.00, SD: 77,450.00). On the contrary, words in semantic-categorical pairs co-occurred in significantly fewer web pages (mean: 25.17, min: 0.00, max: 254.00, SD: 43.38; stimulus set: F (1, 198) = 56.69, MSe = 2.999E9, $p < .001$). The co-occurrence frequency of semantic-categorically related pairs was assessed in four different paratactic constructions for the two nouns of each pair: [$N_{[nom]}$, $N_{[nom]}$], [$N_{[nom]}$ και (‘and’) $N_{[nom]}$], [$DET_{[nom]}$ $N_{[nom]}$, $DET_{[nom]}N_{[nom]}$], [$DET_{[nom]}$ $N_{[nom]}$ και (‘and’) $DET_{[nom]}N_{[nom]}$], where DET was the Modern Greek definite article. The mean number of pages retrieved determined the value of formal co-occurrence of each pair.

[86] Using the ‘advanced search’ utility, web pages containing the queried terms of the ‘exact wording or phrase’ were selected. The search was performed on 15 November 2009.

Word frequency measurements were taken from the ‘‘Hellenic National Corpus’’ (HNC; Hatzigeorgiu *et al.*, 2000), which currently contains more than 47.000.000 words of written texts. The task was comprised of 500 pairs/trials, with 250 trials having a word, and 250 trials having a nonword for a target string. The former 250 trials consisted of 100 unrelated, 100 related word-word pairs, and 50 nonword-word filler pairs. The latter 250 trials included 200 filler nonword-nonword pairs and 50 filler word-nonword pairs. Thus the composition of the stimuli yielded a low relatedness proportion of 1:5, which was far from inducing strategic effects in performance, and invited an equal number of ‘yes’ and ‘no’ responses (e.g. Rogers & Friedman, 2008). The semantic-categorical relatedness type tested here was one of coordinates of a certain semantic category (e.g. αυλή (‘yard’) - χωράφι (‘field’)). Control pairs were constructed by re-arranging the prime and target words of the related pairs into unrelated pairs, and were thus matched with the related ones in prime/target word length and frequency.

The 200 items of the formal-associative and semantic-categorical sets were rotated across four lists; these were generated for counterbalancing their relatedness (they could either appear in an unrelated or related version), and their order of appearance (they could appear in the first half of the session, before TMS, or in the second half, after TMS), as shown in table 6.1 below. Formal-associative and semantic-categorical items were matched for prime and target word length and frequency, and were all counterbalanced across the four lists: (priming type, list, priming type*list: all *F*s, $F < 1$; table B.3 in Appendix B). Similarly, pairs with word targets were matched across two lists (generated for counterbalancing their order of appearance) for prime and target word frequency and length (list: $F < 1$; table B.4 in Appendix B). Pairs with nonword targets were also matched across the two lists for prime word frequency and length, and target word length (list: $F < 1$; table B.5 in Appendix B). Nonword targets however were significantly longer than word-targets (wordness: $F(1, 496) = 14.53$, MSe = 1.49, $p < .001$).

Half	Priming Type	Lists							
		A		B		C		D	
		Unrelated	Related	Unrelated	Related	Unrelated	Related	Unrelated	Related
First Half	Semantic-Categorical	σίδερο-βοδινό (iron-beef)	κλαδί-φύλλο (branch-leaf)	ψάρι-χρυσός (fish-gold)	κερί-λάμπα (candle-lamp)	κλαδί-λάμπα (branch-lamp)	σίδερο-χρυσός (iron-gold)	κερί-φύλλο (candle-leaf)	ψάρι-βοδινό (fish-beef)
	Formal-Associative	διάολος-άχυρα (devil-straws)	αέρας-πανιά (wind-sails)	βελόνα-ποδάρι (needle-leg)	γκάζι-πάτωμα (gas-floor)	αέρας-πάτωμα (wind-floor)	διάολος-ποδάρι (devil-leg)	γκάζι-πανιά (gas-sails)	βελόνα-άχυρα (needle-straws)
Second Half	Semantic-Categorical	ψάρι-χρυσός (fish-gold)	κερί-λάμπα (candle-lamp)	σίδερο-βοδινό (iron-beef)	κλαδί-φύλλο (branch-leaf)	κερί-φύλλο (candle-leaf)	ψάρι-βοδινό (fish-beef)	κλαδί-λάμπα (branch-lamp)	σίδερο-χρυσός (iron-gold)
	Formal-Associative	βελόνα-ποδάρι (needle-leg)	γκάζι-πάτωμα (gas-floor)	διάολος-άχυρα (devil-straws)	αέρας-πανιά (wind-sails)	γκάζι-πανιά (gas-sails)	βελόνα-άχυρα (needle-straws)	αέρας-πάτωμα (wind-floor)	διάολος-ποδάρι (devil-leg)

Table 6.1: Task conditions (experiment 4).

6.2.2. Procedure

Prime and target stimuli were presented in black letters at the center of a white screen of a MacBook using the 'E-Prime 2' (version 2.0; Psychology Software Tools, Inc, 2009) software in a silent, brightly lit room. Each trial consisted of a sequence of three stimuli presented at the same screen location. First a fixation point ('+') was presented for 400 ms, followed by the prime in lower-case letters for 100 ms, which was followed by the target, also in lower-case letters. Unlike the first study, target strings were not presented in upper-case letters. This was because in Modern Greek only lower-case letters receive accent marking (e.g. 'μηχανή', but: 'MHXANH'), and using upper-case letters would thus reduce information for word recognition. Furthermore, the target string in each trial stayed on the screen for 1500 ms, and not indefinitely, as in the first study. This setting was preferred in view of the absence of any organized subject pool of experienced participants in online psycholinguistic, or, a fortiori, TMS experiments. It

was thus deemed that pressures for fast (hence non-strategic) responses should be elevated in the task protocol. As in the first study, no masking mediated the presentation of the prime and the target word, i.e. the SOA value was confounded with that of the prime duration of 100 ms. The duration of the intertrial interval was kept at 600 ms.

6.2.3. TMS Setup

Unlike the other two TMS studies (chapters 5 and 7), where a 'Super Rapid^2' stimulator was used, only a 'Standard Rapid^2' stimulator (The Magstim Company Ltd, Whitland, Dyfed, United Kingdom) was available in the lab. This meant that the Huang *et al.* (2005) continuous theta-burst stimulation (cTBS) protocol (section 4.4.5) could not be applied: While for the ''Rapid^2 option 2'' (''Super Rapid^2''; The Magstim Company Ltd, 2006, p. 49) a 45% of maximum machine output would correspond to a maximum frequency of 63 Hz, for the ''Rapid^2 option 1'', i.e. the ''Standard Rapid^2'' (*ibid*), a 45% maximum machine output corresponds to a maximum frequency of 33 Hz. Even for the 40% stimulus intensity often used in applying the Huang *et al.* (2005) cTBS protocol (N. Muggleton, personal communication, 27 August 2009), the maximum frequency allowable with a Standard Rapid^2 was 37 Hz, far from the 50 Hz required to reproduce the Huang cTBS. For the latter to be applied with a Standard Rapid^2, an output power of 30% at most should be used- this would represent a significant decrease, as compared with both the power used in the experiments reported, and the one often applied (40% maximum machine output).

Hence, the ''modified cTBS'' protocol (Nyffeler *et al.*, 2006) was employed instead, consisting here in a continuous train of 801 pulses delivered in 267 bursts, each burst consisting of 3 pulses at 30 Hz, and 100 ms mediating between bursts. As in all studies, stimulus intensity was maintained at 45% of maximum machine output across participants (section 4.4.6). The modified cTBS protocol has been delivered at 80% (Schindler *et al.*, 2008), 90% (Nyffeler *et al.*, 2008), or 100% of resting motor threshold

(Nyffeler *et al.*, 2009). As discussed above, the 45% of maximum machine output would averagely correspond to over 90% of resting motor threshold (section 4.4.6).

In each session, subjects were required to have a 5-minute break after completion of the TMS train. They completed the first block within 12 minutes, received TMS (scalp coordinate marking and TMS piloting required 5 minutes), and, after the post-TMS break, completed the second block within 12 minutes. There was an optional break time that participants could take in the middle of each block. Each session thus required 35-40 minutes to complete. All subjects tolerated TMS well, reporting mild discomfort on the surrounding area of stimulation. They provided informed consent and received compensation for their participation.

6.2.4. Participants

24 native speakers of Modern Greek (age mean: 26.42 years; min: 18; max: 52; SD: 9.25) were recruited according to the selection criteria in section 4.4.12. They were tested in two different sessions, with 2 to 28 days (mean: 8.79; SD: 6.39 days) mediating between completion of the first and second TMS sessions. One group ('group 1': n = 12) underwent PLCB TMS in the first session and NCBV TMS in the second, and another group ('group 2': n = 12) received NCBV TMS in the first session and PLCB TMS in the second. The two groups did not differ with respect to the age of participants (group 1: 24.33 years; SD: 3.60; group 2: 28.50 years; SD: 12.52) or the time distance mediating between the first and the second TMS session (group 1: 7.92 days; SD 4.76; group 2: 9.67 days; SD: 7.82), and therefore 'group' was not a predictor of either (age: $F < 1$; time distance: $p > .25$).

6.2.5. Design

For the analysis of TMS effects on priming size, the experiment employed a 2 x 2 x 2 within-subjects design. Participants were stimulated in two different CB sites in two different sessions (site: NCBV, PLCB). Their priming sizes were calculated before and after TMS (phase). Sizes were calculated separately for formal-associative and semantic-categorical priming (priming type). For the analysis of TMS effects on sensorimotor performance, another 2 x 2 x 2 analysis was used. Mean latencies and accuracy rates for words and nonwords were calculated separately ('wordness') before and after TMS (phase) on the NCBV or the PLCB (site). Additional analyses were performed including the 'group' variable (1, 2).

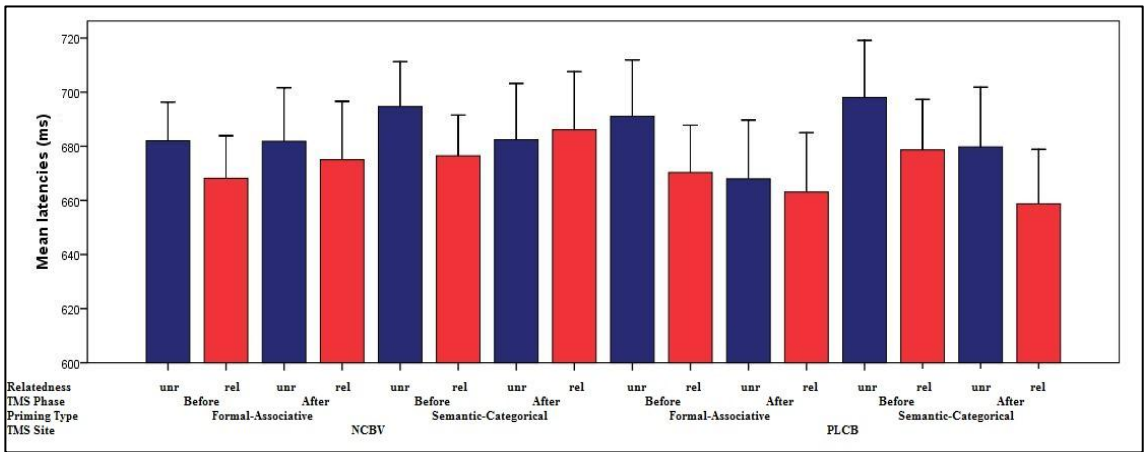
6.2.6. Results

6.2.6.1. Priming Effect

In the first study, because of the small subject group in the TMS experiment (n = 8; section 5.4), a pilot experiment with a larger subject group (n = 18) was run to show that priming occurred in a direct fashion in the form of significantly shorter latencies for related than for unrelated items (section 5.2). On the contrary, the number of participants in this TMS study (n = 24) sufficed to directly show priming. Instead of reporting a pilot study to demonstrate this, an analysis of raw RTs in the main TMS study is provided here, in a 2 x 2 x 2 x 2 design (site, priming type, phase, relatedness). Mean latencies per condition are shown below in plot 6.1 and in table 6.2. The trials excluded according to the criteria in section 4.4.11 amounted to 2.5% of the data.

As demonstrated by **a four-way ANOVA (site, priming type, phase, relatedness)** on lexical decision latencies for unrelated and related pairs, related pairs showed **priming across conditions, as there was a main effect of relatedness (F_1 (1, 20) = 21.33, MSe = 719.51, p < .001; F_2 (1, 192) = 16.07, MSe = 4982.11, p < .001)**,

which was furthermore **not selective for priming type (priming type*relatedness: both *F*s, *F* < 1)**. The rest of the effects were captured in the standard differences-based analysis of priming sizes (section 4.4.11), and are reported below.



Plot 6.1: Mean latencies for unrelated and related pairs per condition for TMS sessions. Error bars represent + 1 SEM (experiment 4).

TMS Site	Priming Type	TMS Phase	Relatedness	Mean latencies (ms)	SD (ms)
NCBV	Formal-Associative	Before	Unrelated	682.07	69.70
			Related	668.16	77.26
		After	Unrelated	681.87	96.89
			Related	675.04	105.69
	Semantic-Categorical	Before	Unrelated	694.66	81.31
			Related	676.49	73.77
		After	Unrelated	682.42	101.94
			Related	686.11	105.22
PLCB	Formal-Associative	Before	Unrelated	691.02	102.18
			Related	670.31	85.79
		After	Unrelated	667.95	106.50
			Related	663.10	107.32
	Semantic-Categorical	Before	Unrelated	698.02	103.39
			Related	678.66	91.63
		After	Unrelated	679.75	108.19
			Related	658.74	98.74

Table 6.2: Mean latencies for unrelated and related pairs per condition for TMS sessions (experiment 4).

6.2.6.2. Effects on Priming Sizes

6.2.6.2.1. Lexical Decision Latencies

The mean priming sizes per condition, calculated as the differences between the mean RTs for unrelated pairs and those for related ones (section 4.4.11), are given below in table 6.3 and plot 6.2. The analysis showed a marginally selective disruption of semantic-categorical priming after NCBV TMS. The same main effects and interactions were observed after including the 'group' variable.

A 3-way ANOVA (site, priming type, phase) showed a weakly selective decrease of semantic-categorical priming sizes after NCBV TMS (site*priming type*phase: $\boldsymbol{F_1}$ **(1, 20) = 2.65, MSe = 1183.53,** $\boldsymbol{p}$ **= .12;** $\boldsymbol{F_2}$ **(1, 192) = 1.14, MSe = 6049.32,** $\boldsymbol{p}$ **= .29)**, along with a weak decrease in priming sizes across sites and priming types (phase: F_1 (1, 20) = 2.80, MSe = 1996.37, p = .11; F_2 (1, 192) = 2.15, MSe = 8136.37, p = .14; site: F_1 (1, 20) = 1.17, MSe = 2418.61, p = .29; F_2 (1, 192) = 1.84, MSe = 5700.22, p = .18; for the rest of Fs, $F < 1$, or $p > .2$).

2-way ANOVAs: NCBV conditions (priming type, phase): An analysis of priming sizes in NCBV conditions, however, showed an only marginal decrease of priming sizes across priming types (phase: F_1 (1, 20) = 2.67, MSe = 1886.08, p = .12; F_2 (1, 192) = 2.91, MSe = 6749.13, p = .09), with no selective decrease for any of the two types **(**priming type*phase: for both Fs, $F < 1$, or $p > .25$)

2-way ANOVAs: PLCB conditions (priming type, phase): No difference in priming sizes was observed after PLCB stimulation (for all Fs, $F < 1$, or $p > .2$).

2-way ANOVAs: Formal-associative priming sizes (site, phase): No significant differences were introduced after TMS of either site on formal-

associative priming sizes (phase: F_1 (1, 20) = 2.35, MSe = 1344.30, p = .14; F_2 < 1; site, site*phase: all Fs, F < 1).

2-way ANOVAs: Semantic-categorical priming sizes (site, phase): An analysis for semantic-categorical priming sizes showed a weakly selective decrease after NCBV TMS (site*phase: F_1 (1, 20) = 2.92, MSe = 1135.52, p = .1; F_2 (1, 96) = 2.11, MSe = 5248.52, p = .15; site: F_1 (1, 20) = 2.79, MSe = 1441.51, p = .11; F_2 (1, 96) = 1.66, MSe = 6711.75, p = .2; phase: both ps, p > .2).

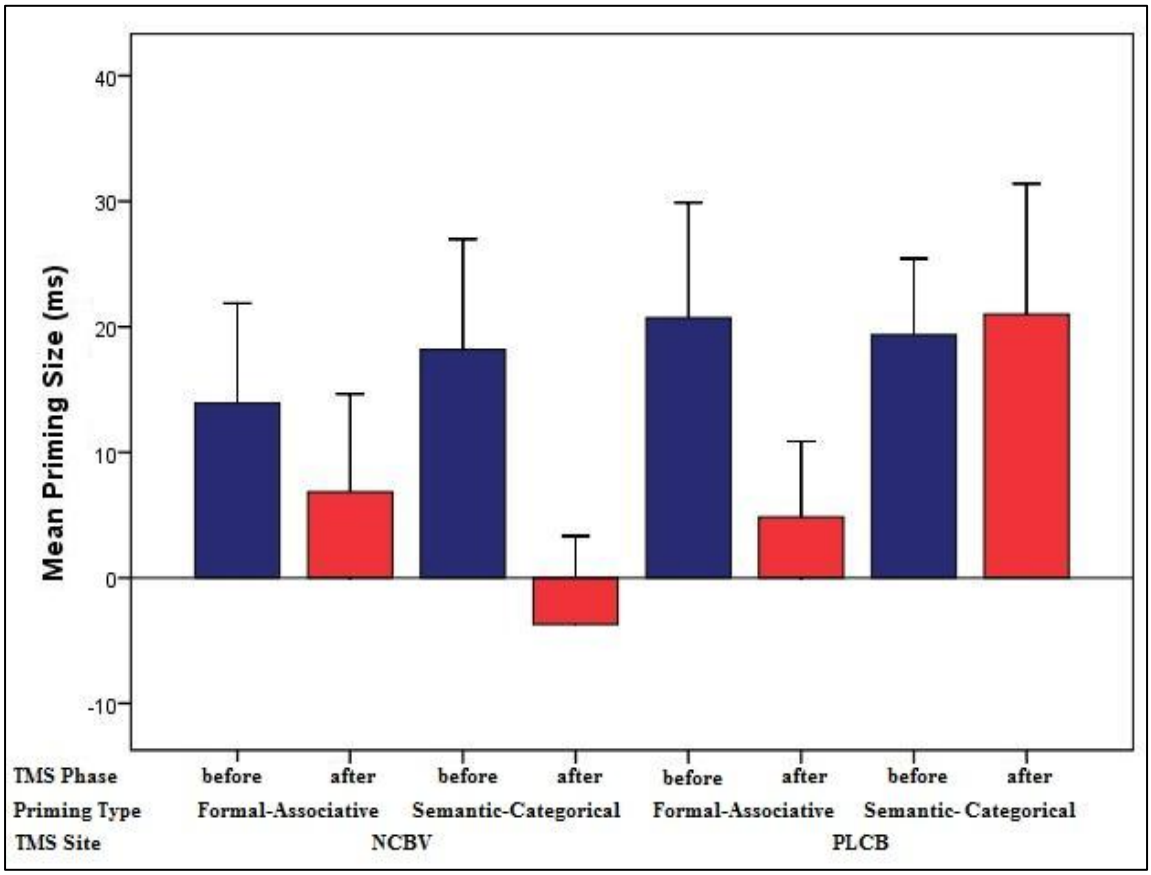
2-way ANOVAs: Before TMS (site, priming type): No significant differences were observed before stimulation of either site (all Fs, F < 1).

2-way ANOVAs: After TMS (site, priming type): An analysis for priming sizes after TMS showed selectively lower semantic-categorical sizes after NCBV TMS, significantly so by subjects **(site*priming type: F_1 (1, 20) = 4.62, MSe = 924.53, p < .05; F_2 (1, 192) = 1.32, MSe = 6399.71, p = .25;** site: F_1 (1, 20) = 1.30, MSe = 2377.10, p = .27; F_2 (1, 192) = 2.21, MSe = 6399.71, p = .14).

1-way ANOVAs (phase): An analysis of semantic-categorical priming sizes for NCBV TMS conditions showed a marginal decrease after TMS (phase: F_1 (1, 20) = 3.35, MSe = 1714.61, p = .08; F_2 (1, 96) = 3.81, MSe = 5954.25, p = .05), while no such changes were observed after **NCBV TMS** for **formal-associative priming** (phase: both Fs, F < 1), or after **PLCB TMS** for either **formal-associative** (phase: F_1 (1, 20) = 1.87, MSe = 1611.97, p = .19; F_2 < 1) or for **semantic-categorical** priming sizes (phase: both Fs, F < 1).

1-way ANOVAs (site): An analysis of semantic-categorical priming sizes between the two sites after TMS showed marginally smaller sizes after NCBV TMS (site: F_1 (1, 20) = 4.01, MSe = 1823.52, p = .06; F_2 (1, 96) = 3.49, MSe = 6366.07, p = .07). No such differences were found in comparisons for **semantic-categorical priming**

sizes before TMS (site: both *F*s, $F < 1$) or for **formal-associative priming sizes before** (site: both *F*s, $F < 1$) or **after TMS** (site: both *F*s, $F < 1$).



Plot 6.2: Mean priming sizes per condition for TMS sessions. Error bars represent + 1 SEM (experiment 4).

TMS Site	Priming Type	TMS Phase	Mean Priming Size (ms)	SD (ms)
NCBV	Formal-Associative	Before	13.91	39.04
		After	6.83	38.30
	Semantic-Categorical	Before	18.18	43.14
		After	-3.69	34.41
PLCB	Formal-Associative	Before	20.71	45.04
		After	4.84	29.57
	Semantic-Categorical	Before	19.36	29.71
		After	21.00	50.95

Table 6.3: Mean priming sizes per condition for TMS sessions (experiment 4).

6.2.6.2.2. Lexical Decision Accuracy

Priming sizes were calculated here on the basis of the difference between the arcsine-transformed accuracy rates for unrelated items and those for related items per condition (section 4.4.11). A **three-way ANOVA (site, priming type, phase)** showed no change in priming sizes after TMS selectively for site or priming type (priming type*phase, site*phase, priming type*site*phase: all *F*s, $F < 1$). The same results were shown in a **four-way ANOVA**, including the 'group' variable.

6.2.6.3. Effects on Sensorimotor Performance

6.2.6.3.1. Lexical Decision Latencies

An analysis of lexical decision latencies for words and nonwords per condition showed that latencies did not become as short after NCBV TMS as after PLCB TMS. Most importantly, NCBV TMS yielded no decrease in RTs selectively when NCBV TMS was applied in the second session of participation, where the stimulus set was presented again to the participants. Mean latencies are shown below in plot 6.3 and table 6.4.

As demonstrated by a **four-way ANOVA (group, site, phase, wordness)** on latencies, the two sites did not differ significantly with respect to RTs (site: $F_1 < 1$; F_2 (1, 496) = 3.38, MSe = 3445.80, $p = .07$), while subject group 1 were significantly by-items faster than subject group 2 across conditions (group: F_1 (1, 20) = 1.30, MSe = 61943.43, $p = .27$; F_2 (1, 496) = 336.20, MSe = 3737.34, $p < .001$). Expectedly, word targets required significantly shorter lexical decision latencies across conditions than nonwords (wordness: F_1 (1, 20) = 36.70, MSe = 3745.23, $p < .001$; F_2 (1, 496) = 111.54, MSe = 24002.60, $p < .001$), and participants performed faster in the second phase across sessions (phase: F_1 (1, 20) = 5.24, MSe = 1852.81, $p < .05$; F_2 (1, 496) = 35.47, MSe = 4570.00, $p < .001$). Most importantly, **latencies did not become any shorter after**

NCBV TMS for group 1, as demonstrated by a strong site*phase*group interaction (F_1 (1, 20) = 12.56, MSe = 566.24, $p < .005$; F_2 (1, 496) = 38.31, MSe = 2893.85, $p < .001$). Furthermore, latencies after NCBV TMS across groups did not become as short as those after PLCB TMS, as demonstrated by the marginal by-subjects and significant by-items site*phase interaction (F_1 (1, 20) = 4.19, MSe = 566.24, $p = .05$; F_2 (1, 496) = 12.66, MSe = 3180.95, $p < .001$). A number of significant interactions are not discussed for the sake of brevity (site*group: F_1 (1, 20) = 73.43, MSe = 1445.86, $p < .001$; F_2 (1, 496) = 586.07, MSe = 3914.12, $p < .001$; wordness*phase: F_1 (1, 20) = 6.14, MSe = 265.57, $p < .05$; F_2 (1, 496) = 5.97, MSe = 4570.00, $p < .05$; wordness*group: F_1 (1, 20) = 5.29, MSe = 3745.23, $p < .05$; F_2 (1, 496) = 82.53, MSe = 3737.30, $p < .001$; phase*group: F_1 (1, 20) = 1.61, MSe = 1852.81, $p = .22$; F_2 (1, 496) = 24.04, MSe = 3539.67, $p < .001$; site*wordness*group: F_1 (1, 20) = 5.38, MSe = 767.17, $p < .05$; F_2 (1, 496) = 15.33, MSe = 3914.12, $p < .001$; for the rest of *F*s, $F < 1$ or $p > .25$).

Due to space restrictions, and since no selective effects were observed for words or nonwords after stimulation of either site (site*phase*wordness, group*site*phase*wordness: all *F*s, $F < 1$), the following analysis was performed collapsing on the 'wordness' variable.

2-way ANOVAs: NCBV TMS conditions (group, phase): An analysis for NCBV TMS conditions demonstrated that latencies became shorter in the second phase across groups, but only significantly so by items (phase: $F_1 < 1$; F_2 (1, 498) = 6.09, MSe = 3200.82, $p < .05$). However, **this reduction was stronger for group 2, as demonstrated by the significant phase*group interaction (F_1 (1, 20) = 6.05, MSe = 798.08, $p < .05$; F_2 (1, 498) = 60.95, MSe = 3200.82, $p < .001$).** As above, group 1 were faster in their lexical decisions than group 2 (group: $F_1 < 1$; F_2 (1, 498) = 22.22, MSe = 3486.73, $p < .001$).

2-way ANOVAs: PLCB TMS conditions (group, phase): On the contrary, for PLCB TMS conditions, no such phase*group interaction was observed (both *F*s, $F < 1$),

with latencies strongly becoming shorter across groups (phase: F_1 (1, 20) = 13.18, MSe = 411.45, $p < .005$; F_2 (1, 498) = 40.40, MSe = 4504.12, $p < .001$). As above, latencies for group 2 were shorter across phases (group: F_1 (1, 20) = 5.37, MSe = 17320.88, $p < .05$; F_2 (1, 498) = 710.84, MSe = 4885.73, $p < .001$).

2-way ANOVAs: Group 1 (site, phase): An analysis of latencies for group 1 showed that **RTs decreased selectively after PLCB, not after NCBV TMS (site*phase: F_1 (1, 10) = 21.50, MSe = 205.73, $p < .005$; F_2 (1, 498) = 46.82, MSe = 3041.43, $p < .001$), and not across sites (phase: $F_1 < 1$; F_2 (1, 498) = 1.38, MSe = 4459.80, $p = .24$).** Expectedly, latencies in the second session of participation (NCBV TMS) were much shorter than those in the first session of participation, i.e. in PLCB TMS conditions (site: F_1 (1, 10) = 23.99, MSe = 992.91, $p < .005$; F_2 (1, 498) = 242.41, MSe = 4081.58, $p < .001$).

2-way ANOVAs: Group 2 (site, phase): On the contrary, an analysis of latencies for group 2 demonstrated but a marginal by-items site*phase interaction ($F_1 < 1$; F_2 (1, 498) = 2.83, MSe = 3087.20, $p = .09$), with latencies becoming significantly shorter only across sites (phase: F_1 (1, 10) = 12.19, MSe = 481.18, $p < .01$; F_2 (1, 498) = 64.14, MSe = 3757.85, $p < .001$). Expectedly, latencies in the second session, i.e. in PLCB TMS conditions were much shorter than those in the first session, i.e. in NCBV TMS conditions (site: F_1 (1, 10) = 64.92, MSe = 452.95, $p < .001$; F_2 (1, 498) = 383.46, MSe = 4342.30, $p < .001$).

2-way ANOVAs: Before TMS (group, site): Unexpectedly, an analysis of latencies before TMS showed a reduction of RTs in the second session of participation selectively for group 1 (site*group: F_1 (1, 20) = 127.00, MSe = 331.13, $p < .001$; F_2 (1, 498) = 497.48, MSe = 3430.68, $p < .001$), and not across sites (site: $F_1 < 1$; F_2 (1, 498) = 1.32, MSe = 3272.89, $p = .25$). As shown above, group 2 was faster than group 1 ($F_1 < 1$; F_2 (1, 498) = 86.23, MSe = 3987.19, $p < .001$).

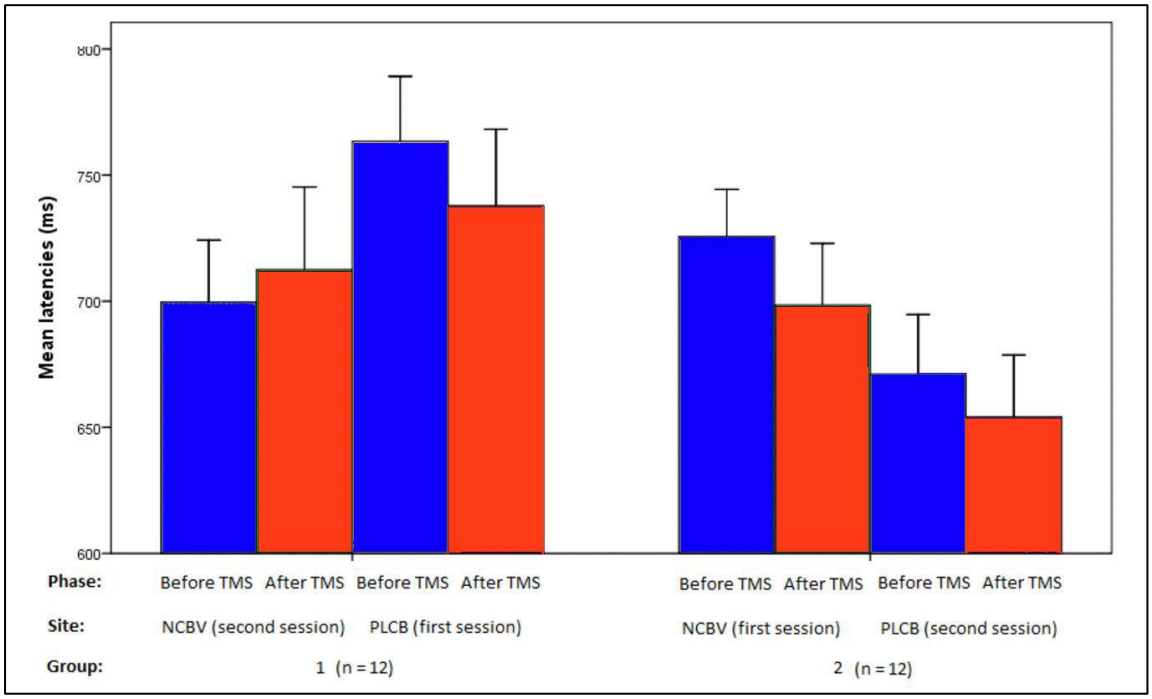
2-way ANOVAs: After TMS (group, site): An analysis of latencies for the post-TMS phase showed that **RTs in the second session reduced across groups significantly by items (site: F_1 (1, 20) = 1.61, MSe = 674.92, p = .22; F_2 (1, 498) = 13.79, MSe = 3452.98, p < .001), but even more after PLCB TMS (site*group: F_1 (1, 20) = 21.61, MSe = 674.92, p < .001; F_2 (1, 498) = 200.26, MSe = 3485.97, p < .001).** As demonstrated above, group 2 were significantly faster by-items than group 1 across sessions (group: F_1 (1, 20) = 1.61, MSe = 674.92, p = .22; F_2 (1, 498) = 251.71, MSe = 3963.93, p < .001).

1-way ANOVAs (site): Comparing RTs in the same phase and for the same group between the two different sessions, significantly shorter latencies were noted in the second session of subjects' participation, i.e. for **group 1 before TMS** (site: F_1 (1, 10) = 85.33, MSe = 285.81, p < .001; F_2 (1, 498) = 260.85, MSe = 3608.38, p < .001), for **group 2 before TMS** (site: F_1 (1, 10) = 47.59, MSe = 376.45, p < .001; F_2 (1, 498) = 248.69, MSe = 3095.20, p < .001), as well as for **group 2 after TMS** (site: F_1 (1, 10) = 27.02, MSe = 437.00, p < .001; F_2 (1, 498) = 162.12, MSe = 3424.31, p < .001). **For group 1 after TMS, latencies in the second session of participation were still shorter than those in the first session, but only significantly by items, and to a smaller extent than that of other such reductions (site: F_1 (1, 10) = 4.23, MSe = 912.84, p = .07; F_2 (1, 498) = 54.22, MSe = 3514.64, p < .001).**

1-way ANOVAs (group): A comparison of latencies in the same phase and session between the two different groups showed similar discrepancies. Expectedly, latencies **before PLCB TMS** for group 1 (first session) were significantly smaller than those for group 2 (second session) in the same phase (F_1 (1, 20) = 6.38, MSe = 8018.54, p < .05; F_2 (1, 498) = 455.79, MSe = 3930.06, p < .001). The same discrepancy was shown for the phase **after PLCB TMS**, yet marginally by subjects and significantly by items (group: F_1 (1, 20) = 4.33, MSe = 9713.78, p = .05; F_2 (1, 498) = 395.88, MSe = 4250.16, p < .001). For the phase **before NCBV TMS**, group 1 (second session) showed significantly by-items shorter latencies than those for group 2 (first session: group: F_1 <

1; F_2 (1, 498) = 74.33, MSe = 3487.81, p < .001). Importantly, **for the phase after NCBV TMS, the latencies of group 1 were significantly by-items shorter than those of group 2, yet to a much smaller extent than that of any of the reductions above (group: F_1 < 1; F_2 (1, 498) = 4.17, MSe = 3199.74, p < .05).**

1-way ANOVAs (phase): A comparison of latencies in the same group and site conditions between the pre- and post-TMS phases showed a similar pattern. For **group 2**, there was a significant reduction of RTs after **NCBV TMS** (F_1 (1, 10) = 9.73, MSe = 458.24, p < .05; F_2 (1, 498) = 46.63, MSe = 3663.20, p < .001), which was also observed after **PLCB TMS**, marginally by subjects and significantly by items (phase: F_1 (1, 10) = 4.51, MSe = 383.45, p = .06; F_2 (1, 498) = 24.82, MSe = 3181.85, p < .001). For **group 1**, the same reduction was observed after **PLCB TMS** (phase: F_1 (1, 10) = 8.90, MSe = 439.44, p < .05; F_2 (1, 498) = 22.50, MSe = 4616.76, p < .001). **However, a much smaller reduction was observed after NCBV TMS for group 1, which was only significant by items, and with a much smaller effect size (phase: F_1 < 1; F_2 (1, 498) = 15.49, MSe = 2884.47, p < .001).**



Plot 6.3: Mean latencies per condition for TMS sessions. Error bars represent + 1 SEM (experiment 4).

Site	Phase	Group	Mean latencies (ms)	SD (ms)
NCBV	**Before**	1	699.64	85.14
		2	725.73	64.57
	After	1	712.50	113.46
		2	698.48	84.76
PLCB	**Before**	1	763.39	89.24
		2	671.09	81.78
	After	1	737.86	105.04
		2	654.12	84.93

Table 6.4: Mean latencies per condition for TMS sessions (experiment 4).

6.2.6.3.2. Lexical Decision Accuracy

Accuracy rates for words and nonwords per condition were arcsine-transformed (section 4.4.11) and subjected to a **four-way ANOVA (group, site, phase, wordness)**. The effects on latencies above were unaccompanied by similar effects on accuracy rates after TMS of a particular site and/ or for a particular group (phase*group, wordness*site*phase*group, wordness*phase*group, site*phase*group, site*phase: for all *F*s, $F < 1$ or $p > .1$). The same pattern was observed in a **three-way ANOVA** excluding the 'group' variable.

6.2.7. Discussion

The TMS experiment above provided evidence for disruptions in the lexical decision process selectively after NCBV TMS: latencies after NCBV TMS did not become as short as after PLCB TMS (section 6.2.7.1). Most importantly, overall lexical decision latencies did not become any shorter after NCBV TMS, but selectively so when NCBV TMS was applied in the second session of participation, where the stimuli were repeated. With respect to TMS effects on priming, a marginally selective disruption was noted for semantic-categorical priming after NCBV TMS (section 6.2.7.2). In view of these marginal effects, a number of limiting conditions are discussed (section 6.2.7.3).

6.2.7.1. Sensorimotor Effects

Similarly to the findings of the first TMS study, the longer latencies observed here after NCBV TMS could be attributed to disruptions in saccadic eye movements or in cognitive-motor associative relations developed between nonwords- 'NO' button, and words- 'YES' button (sections 2.3.4.3 and 5.4.6.2). Reading fluency on words and nonwords was confounded with button presses in the context of the lexical decision task: all participants were right-handed, and were asked to use their right-hand index finger for the word response. However, the disruptions on the speed of lexical decisions were significant across word (right button presses) and nonword targets (left button presses). This corroborates the pattern found in the first TMS study, where accuracy rates dropped after NCBV TMS in the first session of participation across words and nonwords. Thus, a perceptual deficit affecting both words and nonwords would remain at the core of the explanation. Stimulation of a structure directly involved in saccadic movements as the NCBV (see section 5.4.7.1) would better explain such slowing in terms of an overall reading-related disruption.[87] The results above provide the first evidence in the literature for reading-related disruptions induced by NCBV TMS.

6.2.7.2. Cognitive Effects

The results show a reduction in semantic-categorical priming sizes after NCBV TMS, yet only in a marginally selective fashion. This pattern is qualitatively different from that observed in the first TMS study, where NCBV TMS induced an increase in the size of formal-associative priming. However, these findings are far from mutually exclusive. One possibility may be that the disruption induced by NCBV TMS here had an impact on the perception of both prime and target words. Evidence for the latter was

[87] The results do not suffice to discard the cognitive-associative explanation discussed above, yet the oculomotor functions of the NCBV are much better established than the cognitive-associative ones of the NCB in general.

discussed in the previous section. For the former, though, the impact of NCBV TMS may have been greater, given the higher perceptual demands involved: while 1500 ms were given for the perception of and the decision on the target word, only 100 ms were given for the perception of the prime word. Conceivably, then, participants may have failed to perceive the prime word after NCBV TMS. Failing to perceive the prime word in the case of related items would make lexical decision latencies indistinguishable from those for unrelated items.[88] The reason that this was not shown in the case of formal-associative priming would be that a selective, weaker cognitive effect of associative boost occurred, similar to that clearly found in the first study. Such boost might have cancelled out the perceptual disruption induced. In other words, after NCBV TMS, whenever the prime was successfully perceived, the associative boost enabled an acceleration of responses for formal-associatively related pairs. In that case, the lack of a cognitive boost for semantic-categorical priming size would explain its significant decrease after NCBV TMS.

Such explanation would at least require showing that the reading demands in this task were higher than those of the first TMS study, where the prime word was perceived well enough to invite the strong formal-associative priming observed after NCBV TMS. Both prime and target words in each pair here were presented in lower-case letters (section 6.2.2), conceivably thus diminishing the perceptual salience of the prime word. Moreover, the stimuli here appeared in black fonts against a white screen background, in a brightly lit room, whereas in the first study, the letters appeared in green fonts, against a black background, in a dimly lit room. The visual settings of the first study might thus have significantly alleviated any perceptual demands that appeared in this study to a larger extent.

On the other hand, the pattern here may be interpreted on the grounds of a cognitive effect induced by NCBV TMS. Both enhanced and impoverished semantic-

[88] In order to verify this, one would need to involve a separate control task, e.g. testing the perception of words appearing on the screen for only 100 ms (see section 8.2.1 for discussion).

categorical priming have been reported in the literature of schizophrenic populations (for a review, see Picard *et al.*, 2008). Suggestively, schizophrenics have shown associative priming sizes that were significantly larger than those for associatively and semantic-categorically related pairs (Nestor *et al.*, 2006). The first study had also demonstrated a marginal decrease in semantic-categorical priming sizes after NCBV TMS, but sizes in baseline conditions were poorly matched to allow for any conclusions. In this context, the study here would provide suggestive evidence for the other side of the coin shown in the first study. Conceivably, in better experimental settings, a formal-associative priming boost would be accompanied by a semantic-categorical priming reduction after NCBV TMS. In other terms, this can be seen as a cost in processing categorical relations, induced by a change in the trade-off between associative and categorical computations that NCBV TMS might introduce, or by an increased ‘‘associational bias’’ disrupting semantic priming (Nestor *et al.*, 2006) after NCBV TMS. Finally, the weaknesses in constructing the set of formal associates might also be accountable for the lack of an enhancement after NCBV TMS (see next section).

Most importantly, the results above demonstrate a striking lack of decrease of RTs after NCBV TMS, but selectively for the first group, who underwent NCBV TMS in their second session of participation, in which they encountered the same pairs of letter strings presented in the first session. For the explanation of such a selective disruption, then, it would be worth to examine the evidence for NCB involvement in processing repeated linguistic stimuli. However, this CB automatization deficit in task performance (see section 3.5.3) would not pertain here to the task in general, i.e. the routine of reading the prime word, reading the target word, deciding whether the latter is a word, and pressing the appropriate button. Were the disruption to pertain to task performance in general, a similar lack of RT reduction would be expected after NCBV TMS for the second group as well, who received stimulation on that site in the first session of participation, and thus were encountering the letter-string pairs for the first time. As seen above, though, the disruption only appeared for the first group, who were encountering letter-string pairs for the second time. Conceivably, then, the

automatization deficit here pertained in particular to processing items that had been already encountered. A basic factor ensuring shorter lexical decision latencies in the second session of participation may thus be the consolidation of associative memory traces of particular prime-target letter string sequences. In other words, the prime letter string was acquired in the first session as a predictor of the second letter string, independently of any particular relation between the two (e.g. semantic relatedness, phrasal co-occurrence, etc). NCBV TMS may have thus impaired access to and/ or expression of this learned response. This would explain why the same disruption is not found in the second half of the first session of NCBV TMS. If NCB circuitry has stored such prime-target letter strings as sequences of neural events, then, quite plausibly, introduction of noise in such area would disrupt the retrieval of these memory traces, consulting which would ensure faster performance in the corresponding trials. Indeed, in processing repeated stimuli, similar automatization deficits have been observed in working memory and word-generation tasks following CB disruptions (section 2.4.2). Such interpretation is also compatible with recent findings relating CB deficits with compromised cognitive associative learning (Drepper *et al.*, 1999; section 2.3.4.3 here).

The significance of the NCB in linguistic associative learning is in fact in line with the argument in Argyropoulos (2008a, 2009), whereby NCB circuitry provides the foundation for grammaticalization operations. While the latter involve gradual changes manifested at a historical timescale, they are widely held to rely on the automatization of language processing, which results from the repeated articulation-production and perception-comprehension of highly frequent syntactic constructions (sections 3.4 and 3.5 here).

Moreover, the group-specific disruptions in lexical decisions reported here were not shown in Argyropoulos (2010b; chapter 5 here). In that study, the same pairs of letter strings were in fact repeated in a subsequent block for each phase (pre-TMS, post-TMS) in each session. However, no difference was observed between judging repeated letter strings before TMS and judging repeated ones after TMS of any of the two sites

stimulated, either within or between sessions (section 5.4.6.3). A factor to consider would be the significantly smaller number of participants (n = 8) for that study, which conceivably did not suffice to show such a disruption.

6.2.7.3. Limiting Conditions

While the data above showed a rather selective disruption of semantic-categorical priming after NCBV TMS, most of the relevant interactions reached but marginal levels of significance. Furthermore, the fact that the items-based analyses showed stronger and more significant effects both in the priming-related and in the reading-related analyses may suggest that any effects of NCBV TMS occurred only for a subset of participants (and/or that the subject pool size did not suffice). This could be because of the lack of neuronavigational software for the localization of the NCBV. A further limiting factor was, undoubtedly, the criteria used for the composition of the stimulus set. Material from Modern Greek lexical priming studies was not available. Furthermore, the commitment to $N_{[nom]}$-$N_{[nom]}$ sequences discouraged the researcher from undertaking a normative study before the main experiment. Modern Greek N-N sequences of the same case remain much harder to find than those in English- the case factor, being, apparently, a limiting one (see section 6.2.1). Indeed, plenty of words in the formal-associative set were highly idiomatic (e.g. 'σφοντύλι'- idiomatic form of 'σπόνδυλος' ('spine')), often semantically void outwith the idiom, and appear quite infrequently on their own. Furthermore, several words had alternative spellings, according to the simplified spelling preferred in Modern Greek: e.g. 'άσος'- 'άσσος' ('ace'); 'αυτί/αυτιά'-'αφτί/αφτιά' ('ear', 'ears'); 'τραίνο'-'τρένο' ('train'). These issues may have severely compromised automatic lexical access, which is needed to study non-strategic priming effects.

6.3. Conclusion

The second TMS study here showed some rather clear reading-related disruptions selectively induced after NCBV cTBS. Most importantly, latencies after NCBV cTBS became significantly shorter in the first session of participation, but not in the second one. This pattern was discussed as resulting from disruptions in NCB linguistic associative memory access/expression and in the practice-related facilitation in performance supported by NCB circuitry. With respect to the effects of CB TMS on priming sizes, though, results were less rigid. Semantic-categorical priming was disrupted after NCBV TMS in an only marginally selective fashion. Technical limitations in the TMS methodology and stimulus construction were the most important limiting conditions. While further investigation is necessary, the pattern here was discussed in a complementary fashion with that shown in the first TMS study.

Chapter 7

TMS Study 3: Effects of Theta-Burst Stimulation of the Right Neocerebellar Crus I and Vermis on Semantic-Associative and Semantic-Categorical Priming

> ‘‘A calculating machine, an anti-aircraft ‘predictor’, and Kelvin’s tidal predictor all show the same ability. In all of these latter cases, the physical process which it is desired to predict is *imitated* by some mechanical device or model which is cheaper, or quicker, or more convenient in operation. Here we have a very close parallel to our […] stages of reasoning.’’ (Craik, 1943, p. 52; *italics* in the original).

7.1. Introduction

This chapter reports the third TMS study undertaken, part of which is also found in Argyropoulos and Muggleton (2010).

7.1.1. The Previous TMS Studies

The two previous TMS studies (chapters 5 and 6) contrasted noun-to-noun formal-associative with semantic-categorical priming. The first study (chapter 5) showed a selective enhancement of formal-associative priming after TMS of the right neocerebellar vermis (NCBV). The second TMS study (chapter 6), owing, most probably, to technical limitations in TMS application and in stimulus set construction, showed a marginally selective reduction of semantic-categorical priming sizes after

NCBV TMS. Both studies also showed disruptions in the performance in the lexical decision task per se after NCBV TMS; these were explained primarily in terms of disruptions in reading-related processes. The second study also showed disruptions in the practice-related facilitation effects in lexical decision performance after NCBV TMS. Moreover, in those two studies, priming types differed along two axes: linguistic level, i.e. 'semantic vs formal', and type of relation, i.e. 'categorical vs associative' (section 2.3.4.3). This was so because it was difficult to generate a stimulus set for formal-categorical priming without the confounds of repetition or of semantic priming: lexical items that are subordinate to the same morphological category will most probably share a common semantic and formal core (e.g. 'hand', 'handle', 'handling'). Thus, semantic-categorical priming was contrasted with formal-associative priming. According to the theory presented in chapter 2, it was expected that the associative properties would be selectively affected by stimulation of a neocerebellar (NCB) locus. The evidence for preserved semantic-categorical representations in NCB pathology (section 2.4.4.5) supported the idea that semantic-categorical priming would remain unaffected (chapter 5), or would be more susceptible to perceptual disruptions than formal-associative priming (chapter 6) after NCB stimulation. Furthermore, the PLCB site used in the two previous studies served as an adequate control site, either because of its much larger depth, which made it inaccessible to TMS with a conventional flat coil (see section 5.4.7.2.3), or because it did not correspond to a language-related PLCB compartment.

7.1.2. The Present TMS Study

In the present study, the PLCB control site was replaced with another one based on more precise coordinates. Furthermore, the focus was shifted from 'formal-associative' to 'semantic-associative' relations and priming (section 2.3.4.3). In that way, semantic-categorical and semantic-associative priming could be contrasted solely on the basis of the different cognitive mechanisms employed for the two different types of relations (categorical vs associative).

A most characteristic case of semantic-associative priming is that which is induced in cases of thematic (θ)-role-based noun-to-verb priming. In the work of McRae and colleagues, for example, θ-related pairs such as 'sniper-shooting' or 'ball-kicking' received shorter naming latencies compared with unrelated ones (e.g. McRae *et al.*, 2005). Characteristically, in McRae *et al.* (2001), the first study to systematically investigate noun-to-verb priming, it was proposed that event memory is organized in an associative manner, whereby entities and objects activate the event types in which they are typically involved (Lancaster & Barsalou, 1997). In this framework, θ-roles include event (verb)-specific concepts, which are immediately computed online in language processing (McRae *et al.*, 1997) and involve mechanisms of expectancy generation from nouns to verbs or vice versa (McRae *et al.*, 2005).

In chapter 2 (section 2.4.4.5 in particular), these associative, expectancy-based properties of such mechanisms were discussed in the light of the associative, predictive nature of CB computations, along with evidence on the involvement of right PLCB loci in verb-generation tasks (where participants generate appropriate verbs for given noun stimuli; e.g. Petersen *et al.*, 1989; Fiez *et al.*, 1992; Gebhart *et al.*, 2002; Frings *et al.*, 2006). It is thus worth investigating here whether stimulation of NCB loci may selectively interfere with schema-transmission mechanisms from a perception-comprehension perspective. In contrast to the previous studies, then, the one here sought

to examine the involvement of the NCBV and that of a verb generation-related, lateral NCB site in this task (see section 7.4.3 below).

Finally, while the most prominent paradigm employed for capturing such phenomena has been the naming latencies task (e.g. McRae *et al.*, 2005), the lexical decision setup was yet again preferred, given the technical issues that voice recording would present (e.g. prematurely triggering the voice key or not triggering it at all). For example, in the McRae *et al.* (2005) study, these issues were compensated for by the sizeable subject pool used (K. McRae, personal communication, 20 October 2008). In the context of the TMS study, given the very limited size of the subject pool available, sound recording failures would have imposed additional constraints.

7.2. Pilot Experiment 1 (Experiment 5)

In view of the unavailability of a large set of related noun-verb pairs (see below) and a large TMS subject pool, a small number of conditions had to be preserved in order to maintain statistical power; the main TMS experiment was thus decided to contrast no more than two priming types. In order to select a reliable control priming type for the TMS experiment, two different priming types were assessed in a pilot study, along with the third semantic-associative set: formal-associative priming, and semantic-categorical priming.

7.2.1. Stimuli

The session consisted of 480 lexical decision trials, 240 word-nonword filler pairs, and 240 word-word pairs, of which 120 were related. The 1:4 relatedness ratio was quite low, ensuring, along with the low SOA used, minimal interference of strategic effects. Three different noun-verb pair sets were created, with three corresponding priming types. In one set, pairs were 'semantic-categorically' related ('robbery-stealing'). In another set, pairs were 'formal-associatively' related ('baby-sitting'). In the third set, they were 'semantic-associatively' related ('sniper-shooting'). The 80 items of each of the three sets were rotated across 4 lists generated, according to relatedness (appearing half of the time in an unrelated, and half of the time in a related version), and time of appearance in the session (half of the time appearing in the first half of the session, and half of the time in the second half). Thus both subjects-based (F_1) and items-based (F_2) analyses were made possible. Examples of the conditions generated by the rotation of the items of the three sets across the 4 lists created are shown below in table 7.1.

In the 'semantic-categorical' set, the prime-nouns in each pair were virtually synonymous with the verb-targets, e.g. 'robbery-stealing', 'fear-scaring'. Thus, facilitation in processing such related items would be a case of semantic-categorical priming, in which the two terms are synonymous, or, at least, semantic coordinates within the same semantic category. The semantic relatedness of those items was assessed by the 'Wordnet::Similarity' software.[89] In addition, it was ensured that the noun-verb pairs did not immediately co-occur in speech, as assessed in the BNC (written part; Leech, 1992). The four lists contained prime- and target-words of equal mean frequency and length ($F < 1$; table C.2, Appendix C).

The second stimulus set was comprised of semantic-associatively related pairs. The pairs were taken from the stimulus sets used in McRae *et al.* (2005), and Ferretti *et al.* (2001, 2003, 2007). All four major θ-role-based subtypes were employed here, involving agent- (e.g. 'butcher-carving'), patient- (e.g. 'lawn-mowing'), instrument- (e.g. 'scissors-cutting'), and location- (e.g. 'casino-gambling') based priming. In addition, as with the first category, it was ensured that the items did not co-occur in an immediate fashion. Whereas McRae and colleagues (2005) attribute such priming effects to formal (textual) co-occurrence and/or semantic association, the emphasis here was given only to the latter, and, to that end, only pairs with low formal co-occurrence were selected. Thus, pairs like 'wine-tasting', 'ballroom-dancing', for example, that combine formal- and semantic-associative relations, or pairs such as 'toy-playing', that combine formal-associative and semantic-categorical relations, were excluded. Of course, semantic-associative relatedness implies that the forms may co-occur, be it in a non-immediate fashion. However, such loose co-occurrences could easily be found in the categorically related pairs as well. For example, in (1) below, the semantic-associatively related pair 'stripper-entertaining' (1a) can loosely co-occur in speech, like the semantic-categorically related one 'murder-killing' (1b).

[89] As the software cannot assess the semantic relatedness of words of different grammatical categories (here, nouns and verbs), it was ensured that, in each pair, the noun was homonymous with a verb of the same basic semantic properties, e.g. 'release', 'fear'.

(1a) ‘[…] being a male **stripper** isn’t just about taking clothes off, it’s about **entertaining** people […]’

(1b) ‘It would seem that one is guilty of **murder** through **killing** someone by chance […]’

(Google search)

Several more pairs from the item sets of McRae and colleagues were excluded, on different grounds: due to extremely frequent/infrequent words in the BNC (e.g. ‘‘patrolman-searching’’); due to coincidences of the same phoneme in the starting position between prime and target word (e.g. ‘‘crayon-coloring’’), which is avoided in lexical decision setups (Forster & Davis, 1991). Pairs were once again matched across the four lists for length and frequency of prime and target words ($F < 1$; table C.3, Appendix C).

The third stimulus set was comprised of ‘formal-associatively’ related noun-verb pairs: these were predominantly formed on the basis of idiomatic constructions, a great proportion of which included noun-verb backformations. These are much less frequent than those of other grammatical classes, such as the more productive adjective-verb backformations (Bauer, 1983, p. 208), e.g. ‘dry-cleaning’ or ‘long-dividing’. The noun-verb pairs were ensured to be semantic-categorically unrelated. Thus, pairs like ‘sight-reading’ or ‘stock-piling’ were excluded. Noun-verb pairs here were also semantic-associatively (θ-) unrelated. Thus, pairs like ‘ice-skating’ or ‘spoon-feeding’ were not included. It was also ensured that these backformations appeared at least once in the BNC in the N-$V_{[ing]}$ form, and at least once in a non-periphrastic form, e.g. ‘babysitting’, but also ‘baby-sitting’. Once again, the four lists involved prime- and target-words of equal frequency and length (all *F*s, $F < 1$; table C.4, appendix C).

Given the limited number of pairs that could be selected form the verb-generation norms of McRae and colleagues, and the limited number of idiomatically co-

occurring noun-verb sequences, no matching between the different item sets was attempted for prime/target word frequency/length: the three sets differed on prime word frequency (stimulus set: F (2, 237) = 17.47, MSe = 1929.38, $p < .001$), prime word length (F (2, 237) = 19.59, MSe = 3.52, $p < .001$), target word frequency (F (2, 237) = 5.83, MSe = 784.96, $p < .005$), and target word length (F (2, 237) = 8.79, MSe = 1.80, $p < .001$). However, the four lists across which the three sets were rotated were matched for the above measures, separately for each set. Crucially for stimulus construction here, the semantic-categorically related pairs involved semantically more similar nouns and verbs than those of the other two sets according to the 'Wordnet::Similarity' ratings (stimulus set: F (2, 237) = 388.3, MSe = 0.02, $p < .001$; tables C2-4, Appendix C). On the other hand, the formal-associatively related items involved nouns and verbs that co-occurred in speech more frequently than the pairs of the other two sets (stimulus set: F (2, 237) = 45.12, MSe = 0.07, $p < .001$; tables C2-4, Appendix C) .[90] Furthermore, the word-nonword filler pairs were matched for prime word frequency and length, and target nonword length across the two lists generated for counterbalancing their order of appearance (rotation list: all Fs, $F < 1$; table C5, Appendix C).

[90] A disadvantage in the design was, of course, that there was no normative study or quantitative measure available to assess any θ-role-based properties in the semantic-categorical and formal-associative group.

PHASE	Priming Type	Lists							
		A		B		C		D	
		Unr	Rel	Unr	Rel	Unr	Rel	Unr	Rel
HALF 1	Semantic-Categorical	sack-allowing	damage-hurting	permission-firing	guess-predicting	damage-predicting	sack-firing	guess-hurting	permission-allowing
	Formal-Associative	finger-bathing	dynamite-fishing	sun-licking	mud-wrestling	dynamite-wrestling	finger-licking	mud-fishing	sun-bathing
	Semantic-Associative	carpenter-brushing	merchant-selling	teeth-hammering	nun-praying	merchant-praying	carpenter-hammering	nun-selling	teeth-brushing
HALF 2	Semantic-Categorical	permission-firing	guess-predicting	sack-allowing	damage-hurting	guess-hurting	permission-allowing	damage-predicting	sack-firing
	Formal-Associative	sun-licking	mud-wrestling	finger-bathing	dynamite-fishing	mud-fishing	sun-bathing	dynamite-wrestling	finger-licking
	Semantic-Associative	teeth-hammering	nun-praying	carpenter-brushing	merchant-selling	nun-selling	teeth-brushing	merchant-praying	carpenter-hammering

Table 7.1: Task conditions (experiment 5).

7.2.2. Procedure

Prime and target stimuli were presented in black letters at the center of a white screen of a MacBook using the 'DirectRT' software (Jarvis, 2008). Each trial consisted of a sequence of three stimuli presented at the same screen location. First, a fixation point ('+') was presented for 250 ms, followed by the prime for 150 ms, which was followed by the target letter string. The targets remained on the screen until participants responded. No masking mediated the presentation of the prime and the target, i.e. the SOA value was confounded with that of the prime duration of 150 ms (SOAs of 200ms or shorter guarantee automatic processes in priming; e.g. Perea & Rosa, 2002).[91] The intertrial interval lasted 500 ms.

[91] The slightly longer SOA used in the present study (150 ms instead of the 100 ms of the last two studies) was motivated by the longer SOAs used in McRae *et al.* (2005).

Both primes and targets were always presented in lower-case letters. Participants were instructed to focus on the fixation point and to read the first letter string and respond only to the second one. The order of stimulus presentation was randomized for each subject. Subjects were instructed to press one of two buttons on the keyboard (‘j’ for yes and ‘f’ for no) to indicate whether the upper-case letter string was an English word or not, as rapidly and accurately as possible. They used their dominant right-hand index finger for the word responses. When the subject responded, the target disappeared from the screen. Each subject received a total of 20 practice trials prior to the 500 experimental trials. Participants were tested individually in a silent, brightly lit room. They received written instructions explaining the task. All subjects provided informed consent and received compensation for each session.

7.2.3. Participants

12 participants were recruited according to the selection criteria in section 4.4.12; they were rotated across the 4 different lists created.

7.2.4. Design

A 2 x 3 within-subjects and -items design was employed. Prime and target words in a pair could be related or unrelated (relatedness), and they could be related in three different ways (priming type): semantic-categorically (‘robbery-stealing’), semantic-associatively (‘sniper-shooting’), formal-associatively (‘baby-sitting’). The dependent measure was latencies.

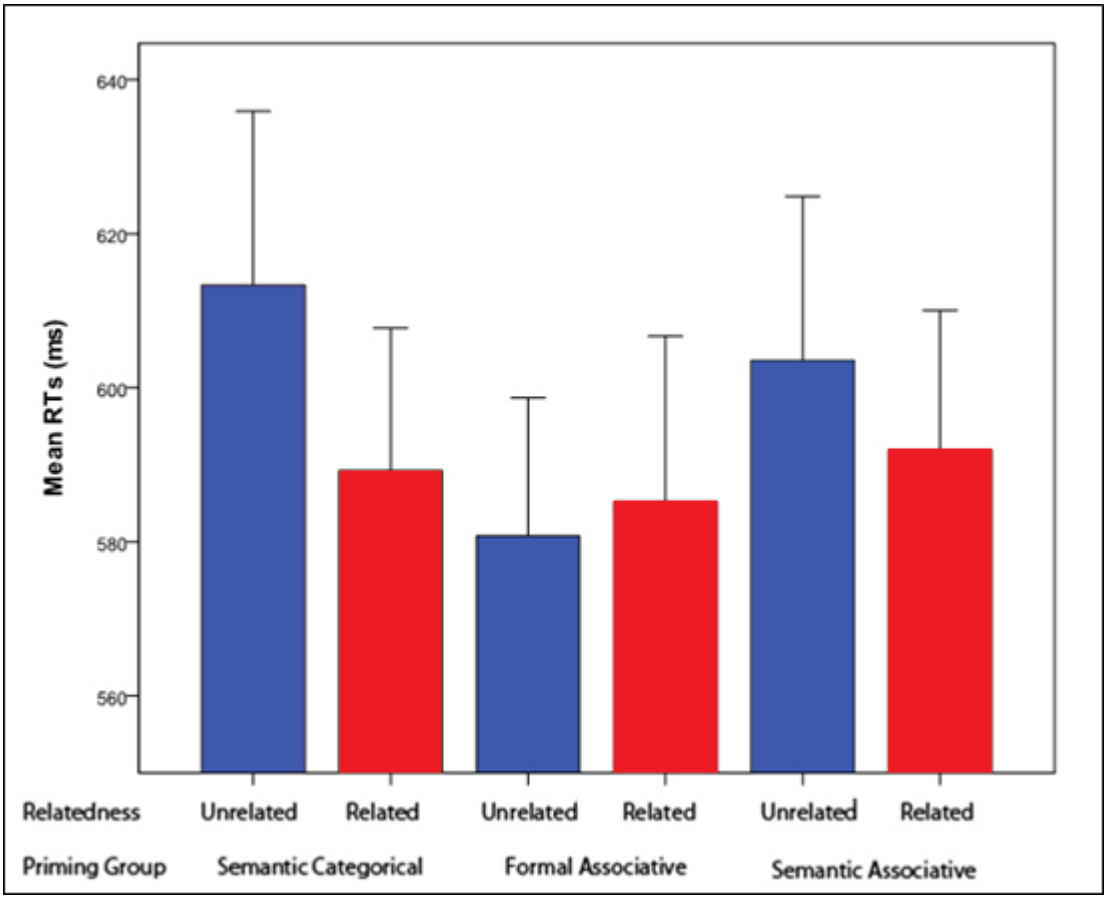
7.2.5. Results

Mean latencies per condition are shown in table 7.2 and plot 7.1 below. The trials excluded according to the criteria in section 4.4.11 accounted for 2.4% of the data. Three trials were also excluded from the data for different reasons,[92] amounting to 1.3% of the data.

A **2-way ANOVA (priming type, relatedness)** demonstrated that, since the three priming sets were not matched for word length or frequency, latencies expectedly differed with respect to priming type (F_1 (2, 16) = 5.97, MSe = 384.56, $p < .05$; F_2 (2, 228) = 6.33, MSe = 3448.50, $p < .005$). Crucially, however, **related items received shorter latencies across priming types (relatedness: F_1 (1, 8) = 8.88, MSe = 219.67, $p < .05$; F_2 (1, 228) = 4.89, MSe = 2795.02, $p < .05$). The three groups, however, differed significantly (by items) in priming strength (priming type*relatedness: F_1 (2, 16) = 2.68, MSe = 455.84, $p = .1$; F_2 (2, 228) = 3.44, MSe = 2730.58, $p < .05$)**, with the semantic-categorical set showing the strongest relatedness effect, and the formal-associative set involving no priming (see below).

1-way ANOVAs (relatedness): Separately for each group, a comparison of latencies for unrelated and related pairs demonstrated strong **semantic-categorical priming** both by subjects and items **(relatedness: F_1 (1, 8) = 10.3, MSe = 335.92, $p < .05$; F_2 (1, 76) = 10.27, MSe = 2854.40, $p < .005$)**. **Semantic-associative priming** was **found only by subjects (relatedness: F_1 (1, 8) = 6.1, MSe = 132.90, $p < .05$; F_2 (1, 76) = 1, MSe = 2696.91, $p > .3$)**. However, no **formal-associative priming** was shown (relatedness: both *F*s, $F < 1$).

[92] Three trials were excluded from the analysis: 'expulsion/guess-ousting', because the target verb was not recognized as a word by more than half of the participants; 'celebrity/egg–recognising', 'stress/collection– emphasizing, for not consistently following the UK or the US English spelling of the –ising/-izing morpheme.



Plot 7.1: Mean latencies for unrelated and related pairs for pilot group per condition. Error Bars represent + 1 SEM (experiment 5).

Priming Type	Relatedness	Mean RTs (ms)	SD (ms)
Semantic-Associative	**Unrelated**	603.55	73.65
	Related	591.92	62.76
Semantic-Categorical	**Unrelated**	613.33	78.13
	Related	589.29	64.00
Formal-Associative	**Unrelated**	580.80	62.03
	Related	585.23	74.44

Table 7.2: Mean latencies for unrelated and related pairs for pilot group per condition (experiment 5).

7.2.6. Discussion

The results above demonstrated that priming occurred only for the semantic-categorical and the semantic-associative sets. The latter showed significant priming only by subjects. This would mean that only certain items in that set involved priming, or that the number of subjects did not suffice to stabilize the variance for the analysis by items. As no priming occurred for the formal-associative types, this set was not included in the main TMS experiment.

7.3. Pilot Experiment 2 (Experiment 6)

In order to determine whether priming still occurred with only the two sets that showed priming (semantic-associative and semantic-categorical), a second pilot experiment was run, with additional items for each of these two sets.

7.3.1. Stimuli

Each session consisted of a total of 500 trials, consisting of 250 filler word-nonword pairs, and 250 word-word pairs; the latter consisted of 100 related, 100 unrelated, and 50 unrelated filler word-word pairs. Thus the relatedness ratio was even lower than that of the first pilot experiment (1:4; section 7.2.1; here: 1:5), reaching that of the previous TMS studies (sections 5.2.1 and 6.2.1). Once again, prime and target word length and frequencies were kept steady across the different blocks, for both item sets. Filler items were matched for prime length, prime word frequency and target word length across the two lists generated for counterbalancing their order of appearance ($F < 1$). No significant differences were found in prime/target word length or prime word frequency between filler (word-nonword) and test (control and target; word-word) items (all *F*s, $F < 1$; table D.4, Appendix D). Categorically related noun-verb pairs were by far more semantically similar as assessed by 'Wordnet::Similarity' (F (1, 198) = 553.5, MSe = 0.03, $p < .001$; mean = 0.73, SD: 0.25, min: 0.5, max: 1; table D.2, Appendix D) than associatively related ones (mean = 0.12, SD = 0.06, min: 0.05 max: 0.33; table D.3, Appendix D). However, co-occurrence strength was kept low and matched for the two sets (mean: 0.07, SD: 0.27; group: $F < 1$; tables D3-4), so that the two priming types would maximally differ along the axis of 'categorical vs associative'. In the same fashion with the experiment above, both categorical and associative sets were balanced across the four lists in prime/target word length/frequency (list: for categorical *F*s, $F < 1$;

for associative *F*s: $F < 1$; tables D3-4). Examples of the conditions created by the rotation of items across the 4 lists are shown below in table 7.3.[93]

PHASE	Priming Type	Lists							
		A		B		C		D	
		Unrelated	Related	Unrelated	Related	Unrelated	Related	Unrelated	Related
HALF 1	Categorical	sack-allowing	damage-hurting	permission-firing	guess-predicting	damage-predicting	sack-firing	guess-hurting	permission-allowing
	Associative	carpenter-brushing	merchant-selling	teeth-hammering	nun-praying	merchant-praying	carpenter-hammering	nun-selling	teeth-brushing
HALF 2	Categorical	permission-firing	guess-predicting	sack-allowing	damage-hurting	guess-hurting	permission-allowing	damage-predicting	sack-firing
	Associative	teeth-hammering	nun-praying	carpenter-brushing	merchant-selling	nun-selling	teeth-brushing	merchant-praying	carpenter-hammering

Table 7.3: Task conditions (experiments 6 and 7).

7.3.2. Procedure

The same procedure as that for the previous pilot experiment (section 7.2.2) was followed. In order to encourage fast lexical decisions, a message appeared on the screen ('Please try to respond faster!') after each trial in which the subject would respond after 1500 ms.

7.3.3. Participants

40 participants were recruited according to the selection criteria in section 4.4.12; they were rotated across the 4 stimulus lists.

[93] Since there was no a priori hypothesis about effects on the particular strengths of different θ-role relations (agent, patient, instrument, location), and given that the experiment examined θ-role priming in general as opposed to semantic-categorical priming, pairs of the four basic θ-role relations were evenly distributed across the four lists. Each list contained 9 agent-, 5 instrument-, 4 location-, and 7 patient-verb pairs.

7.3.4. Design

The present experiment employed a 2 x 2 x 2 within-subjects design. The pairs were either related or unrelated (relatedness), and they could be related in two different ways: associatively or categorically (priming type). Latencies were measured in the first and second phase of the session (phase). This was important, in order to capture any changes in priming sizes occurring in task performance, which should be taken into account for the main TMS experiment.

7.3.5. Results

The mean latencies for related and unrelated items per priming type and per phase are shown in plot 7.2, and table 7.4 below. The trials deleted according to the criteria in section 4.4.11 amounted to 1.9% of the data. 6 trials were also excluded from the analysis, [94] amounting to 1.2% of the data.

A **3-way ANOVA (phase, priming type, relatedness)** showed, expectedly, that latencies became significantly shorter in the second phase of the session (phase: F_1 (1, 36) = 19.94, MSe = 2112.30, $p < .001$; F_2 (1, 192) = 123.45, MSe = 1075.51, $p < .001$). No difference was observed in the latencies between priming types (priming type: both *F*s, $F < 1$; rest of *F*s, $F < 1$). Crucially, however, **priming occurred selectively in the second phase of the session (phase*relatedness: F_1 (1, 36) = 5.32, MSe = 219.92, $p < .05$; F_2 (1, 192) = 5.24, MSe = 1084.16, $p < .05$**; see below), and not across phases (relatedness: both *F*s, $F < 1$).

[94]The trial 'experience/expulsion-ousting' was deleted, as over half of the subjects did not recognize 'ousting' as a word; Other trials were deleted because of inconsistencies in the ising/izing spelling: 'horizon– authorizing', 'highway/celebrity-recognising', stress/ablation-emphasizing; The trial 'prison-punished' was also deleted, because it accidentally had the target word in the '–ed' form. Finally, the trial 'babysitter/binoculars-punishing' was deleted, because of the accidental repetition of the word 'punish'.

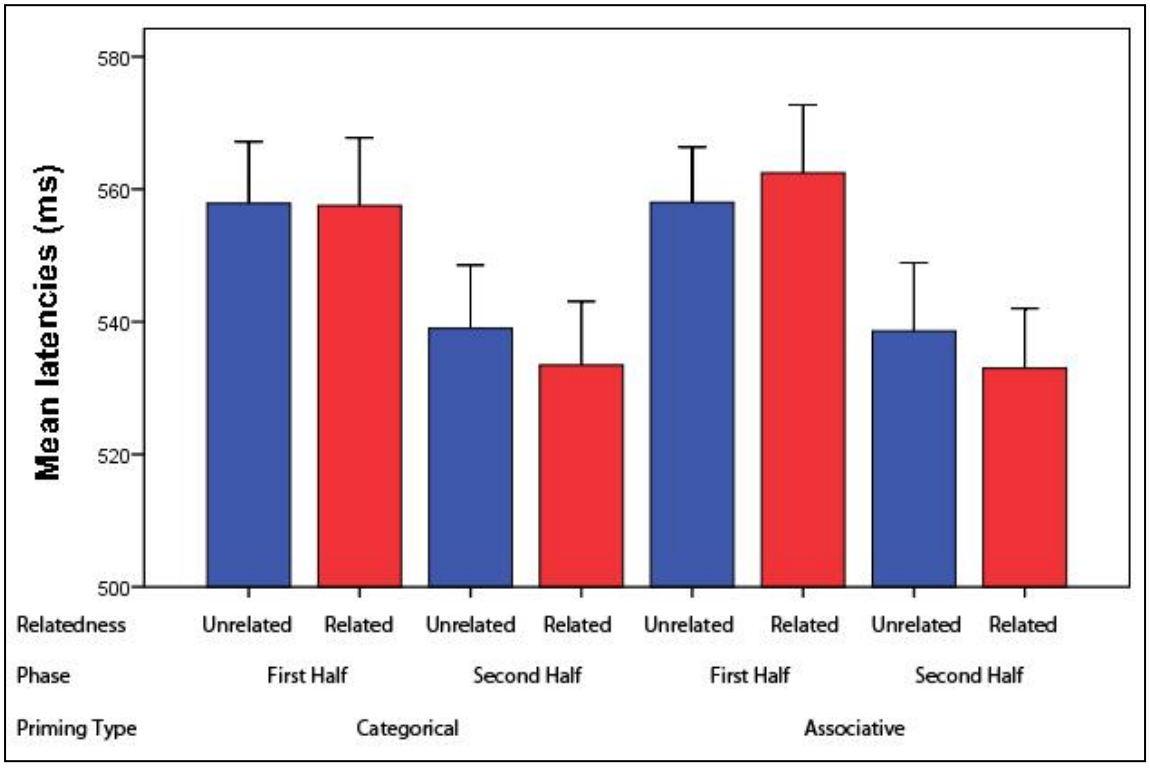
2-way ANOVAs (priming type, relatedness): first phase: An analysis of latencies in the first half showed no priming across priming types or selectively for one (all *F*s, $F < 1$).

2-way ANOVAs (priming type, relatedness): second phase: An analysis for the second phase showed **significant priming (relatedness: F_1 (1, 36) = 4.36, MSe = 286.96, $p < .05$; F_2 (1, 192) = 5.78, MSe = 1243.09, $p < .05$)**, independently of priming type (priming type*relatedness: both *F*s, $F < 1$).No difference in latencies between the two stimulus sets was found in either analysis (priming type: all *F*s, $F < 1$).

An additional analysis was performed, comparing the 4 subtypes of semantic-associative priming, based on the four different θ-roles involved. The variables involved were thus θ-role (4; agent, instrument, location, patient), phase (2; first or second half of the session), relatedness (2; unrelated or related items). A three-way ANOVA demonstrated that the acquisition of priming size in the second phase was not selective for a θ-role subset (for θ-role*phase*relatedness: both *p*s, $p > .35$).

Priming Type	Phase	Relatedness	Mean latencies (ms)	SD (ms)
Categorical	**First Half**	**Unrelated**	557.86	59.02
		Related	557.53	64.73
	Second Half	**Unrelated**	539.01	60.37
		Related	533.45	60.98
Associative	**First Half**	**Unrelated**	558.00	52.76
		Related	562.44	64.98
	Second Half	**Unrelated**	538.61	65.10
		Related	532.98	57.06

Table 7.4: Mean latencies for unrelated and related pairs for pilot group per condition (experiment 6).



Plot 7.2: Mean latencies for unrelated and related pairs for pilot group per condition. Error bars represent + 1 SEM (experiment 6).

7.3.6. Discussion

Quite unexpectedly, the pilot experiment here did not show an overall priming effect, much unlike the first one. Priming occurred only in the second phase across priming types. This could suggest that strategic processes were accountable for. For the main TMS experiment, then, the main hypothesis was that any differences induced by TMS would be manifested with a significantly smaller or larger increase in priming size in the second phase, selectively for associative priming.

7.4. TMS Experiment (Experiment 7)

The TMS experiment was conducted in the Institute of Cognitive Neuroscience (University College, London), hosted by Dr. Neil Muggleton (March-May 2010).

7.4.1. Stimuli

The same stimulus lists were used as those in the second pilot experiment (section 7.3.1).

7.4.2. Procedure

The same procedure was followed as that in the second pilot experiment (section 7.3.2). Upon completion of the first half of the session, the TMS session was applied.

7.4.3. TMS Setup

The Huang *et al.* (2005) continuous theta-burst stimulation protocol (cTBS) was employed, as in the first study (see sections 4.4; 5.4.3). Ethical approval was granted by the local ethics committee. All subjects provided their informed consent and received compensation for their participation.

As in the other two studies, the right NCBV was aimed by placing the coil 1 cm below the inion and 1 cm to the right (e.g. Ohtsuka & Enoki, 1998; sections 5.4.3 and 6.2.3). Instead of the PLCB site used in the two previous studies (1 cm below the inion, 4.5 cm laterally to the right), a yet more lateral right CB hemispheric site was selected. Based on the arguable similarity of the psycholinguistic mechanisms involved in verb-generation and θ-role-based noun-to-verb priming, brain loci involved in the former

become of significant relevance for the latter. In the fMRI study of Frings *et al.* (2006; figure 7.1 below), a comparison of the task conditions "verb generation in inner speech" and "verb reading in inner speech" as a measure of verb generation showed CB activations in lobule HVI (x = 30 mm, y = −56 mm, z = −34 mm, T-value 6.05) and Crus I of the right PLCB (x = 48 mm, y = −60 mm, z = −30 mm, T-value 12.36). The latter locus was selected, given the much larger T-value yielded, as well as the larger depth of the former, which exceeded 3 cm (see section 4.4.9). An advantage, of course, is that, as shown in figure 7.1, the two areas seem to form a rather continuous activated region. This offers a much broader region possibly involved in the task, thus alleviating the costs of not using a neuronavigational system. The Talairach coordinates were registered and converted into scalp coordinates using the Brainsight™ TMS-MRI co-registration system (Rogue Research, Montreal, Canada) in one volunteer, and were found to correspond to 10 cm laterally to the right from the inion. [95] The 10 cm from the inion to the right were also found to correspond to the area activated in a stem completion task (x = 41 mm, y = −55 mm, z = −18 mm) for both overt and covert stem completion (Palmer *et al.*, 2001).

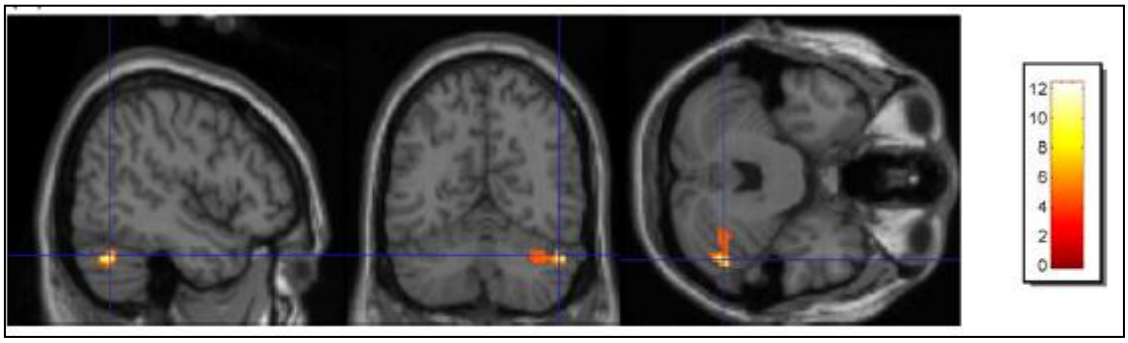


Figure 7.1: Localization of the right NCB Crus I site (lateral-most, yellow-colored site). The two areas shown here are the products of the comparison of the conditions "verb generation in inner speech" and "verb reading in inner speech" as a measure of verb generation. The grid on the right represents T-values corresponding to the activity maximum of each CB area. Figure adapted from Frings *et al.* (2006, p. 21) with permission © 2006 Elsevier.

[95] Thanks to Dr. Neil Muggleton for volunteering and helping in the registration process.

All participants tolerated TMS well, reporting mild discomfort, especially during stimulation of the lateral site, which induced considerably more muscle twitches in the right-side face and neck muscles. The session lasted approximately 30-35 minutes: 12 minutes for the first block, 12 minutes for the second block, and 7 minutes for TMS preparation, application, and for the post-TMS break.

7.4.4. Participants

24 subjects participated in the main TMS study: 8 underwent stimulation of the Crus I (mean age: 24.5, SD: 6.19), 8 underwent stimulation of the NCBV (mean age: 21.63, SD: 4.34), and 8 received a 5-7-minute break without TMS (mean age: 22.50, SD: 8.38). Groups did not differ in mean age ($F < 1$).

7.4.5. Design

The TMS experiment here differed from the two previous ones in that a between-subjects design was preferred. This was owed to the fact that priming occurred selectively in the second phase, and there thus was an unexpected increase in priming size (section 7.3.5.1.1 above). This was not the case in the pilot experiments of the first TMS study (sections 5.2.5.1 and 5.3.5.1) or in the second TMS study in control PLCB TMS conditions. Thus, any change in priming sizes because of TMS had to be directly compared with the changes that occurred in the no TMS group (section 7.3.6). Otherwise, an increase in priming size after TMS would have probably been interpreted as an expected increase, thus suggesting that TMS had no effect.[96] Furthermore, there were no real 'control site' conditions here. Crus I TMS was expected to selectively affect semantic-associative priming; however, it was an open question whether NCBV

[96] Similarly, without direct comparisons with a no TMS group, it would be hard to show whether enhancements selective for priming type should be attributed to poor enhancement for one type or to strong enhancement for the other.

TMS would induce an associative boost, as shown in the first study (chapter 5), or a semantic-categorical disruption, as suggested in the second one (chapter 6).

Hence, for the examination of TMS effects on priming sizes (section 4.4.11), the design here included subject group (3; Crus I TMS, NCBV TMS, no TMS) as a between-subjects factor, as well as priming type (2; associative, categorical) and phase (2; before, after TMS) as within-subjects factors.

For the examination of TMS effects on sensorimotor performance, a similar 3 x 2 x 2 design was employed. Latencies and accuracy rates were analyzed for each of the three groups (group; 3), for the first and second halves of the session (half; 2), for words and nonwords (wordness; 2).

7.4.6. Results

The analysis was performed as in the previous experiment (section 7.3.5). The 6 trials excluded represented 1.2% of the data. The trials excluded by RT trimming represented 1.4% of the data. [97]

7.4.6.1. Effects on Priming Sizes

7.4.6.1.1. Lexical Decision Latencies

7.4.6.1.1.1. Comparison with a no TMS Group

A comparison of priming sizes for the three groups showed a rather selective enhancement of associative priming after stimulation of the Crus I that was much stronger than that for the other two groups. It also showed a rather selective absence of such enhancements for the NCBV group. Mean sizes per group and priming type are

[97] The application of planned comparisons or post-hoc tests did not compromise the significance or the size of the effects reported.

shown below in table 7.5 and plot 7.3. The analysis will start with comparisons among the three groups (three-, two-, and one-way ANOVAs), and will proceed to comparisons between groups, terminating with the ANOVAs for each separate group.

3 groups: A **three-way ANOVA (group, priming type, phase)** showed a difference in overall priming sizes among the three groups, marginally by subjects, and significantly by items (group: F_1 (2, 12) = 2.96, MSe = 412.29, p = .09; F_2 (2, 384) = 3.41, MSe = 7039.81, p < .05). Given the small subject pool size, priming sizes were not matched across groups in baseline conditions. **There was also a significant by-items group*phase interaction, due to the selective lack of increase of priming sizes in the second phase for the NCBV group (F_1 (2, 12) = 1.76, MSe = 652.32, p > .2; F_2 (2, 384) = 3.90, MSe = 6579.68, p < .05)**. The analysis also showed **a priming type*phase*group interaction (F_1 (2, 12) = 5.25, MSe = 561.70, p < .05; F_2 (2, 384) = 3.90, MSe = 6766.75, p < .05;** rest of *p*s, p > .2), owed, to a large extent, to the strong associative priming observed in the second phase for the Crus I group.

3 groups: two-way ANOVAs: Associative priming (group, phase): An analysis of associative priming sizes in particular demonstrated **an increase in the second phase that is selective for the Crus I group (phase*group: F_1 (2, 12) = 6.50, MSe = 449.78, p < .05; F_2 (2, 192) = 4.32, MSe = 7020.47, p < .05)**, and not found across groups (phase: F_1 (2, 12) = 2.79, MSe = 449.78, p = .12; F_2 (1, 192) = 1.02, MSe = 7452.81, p > .3; group: F_1 (2, 12) = 1.1, MSe = 301.06, p > .35; F_2 (2, 192) = 1.68, MSe = 7457.63, p = .19).

3 groups: two-way ANOVAs: Categorical priming (group, phase): An analysis for categorical priming sizes showed a significant by-items phase*group interaction (F_1 (2, 12) = 1.53, MSe = 764.24, p > .25; F_2 (2, 192) = 3.41, MSe = 6138.90, p < .05), due to the selective increase for the no TMS group in the second phase. This change was not found across groups (phase: both *F*s, F < 1; group: F_1 (2, 12) = 1.85, MSe = 601.31, p = .2; F_2 (2, 192) = 1.83, MSe = 6621.99, p = .16).

3 groups: two-way ANOVAs: First half (group, priming type): Comparing priming sizes in the first half showed no significant differences across priming types (group: both *p*s, $p > .35$). There was an only marginal by-subjects group*priming type interaction, due to the high value of categorical priming for the Crus I group in baseline conditions ($F_1(2, 12) = 3.61$, MSe = 577.33, $p = .06$; $F_2(2, 384) = 1.7$, MSe = 6714.01, $p > .2$).

3 groups: two-way ANOVAs: Second half (group, priming type): Sadly, however, a comparison of priming sizes in the second half showed **a weak non-significant group*priming type interaction ($F_1(2, 12) = 2.25$, MSe = 474.46, $p = .15$; $F_2(2, 384) = 2.26$, MSe = 6905.48, $p = .11$)**, despite the large associative priming size for the Crus I group. In a stronger fashion, the **NCBV group demonstrated the lowest priming sizes** in the second phase across priming types, significantly by items, and marginally by subjects **(group: $F_1(2, 12) = 2.91$, MSe = 667.76, $p = .09$; $F_2(2, 384) = 6.19$, MSe = 6905.48, $p < .005$)**.

3 groups: one-way ANOVAs (group): A comparison of the three groups per condition demonstrated marginally stronger **categorical priming** in the **first phase** for the Crus I group (group: $F_1(2, 12) = 3.11$, MSe = 548.59, $p = .08$; $F_2(2, 192) = 2.51$, MSe = 6035.89, $p = .08$). As seen in plot 7.3, the no TMS group involved the largest **categorical priming size** in the **second phase**, but only marginally by items ($F_1 < 1$; $F_2(2, 192) = 2.66$, MSe = 6725.00, $p = .07$). The groups did not differ significantly for **associative priming** size in the **first phase** (group: $F_1(2, 12) = 1.88$, MSe = 425.58, $p = .19$; $F_2 < 1$). Importantly, however, **the Crus I group involved, in a clear fashion, the largest associative priming size in the second phase ($F_1(2, 12) = 7.50$, MSe = 325.26, $p < .01$; $F_2(2, 192) = 5.70$, MSe = 7085.97, $p < .005$)**.

The analysis now turns to a comparison of priming sizes **between the Crus I and the no TMS groups**.

2 groups (Crus I vs no TMS): A **three-way ANOVA (group, priming type, phase)** demonstrated a significant by-items increase in priming sizes overall in the second phase across groups (phase: F_1 (1, 8) = 3.56, MSe = 552.81, p = .1; F_2 (1, 192) = 6.80, MSe = 5989.44, p < .05). The Crus I group involved marginally larger by items priming sizes overall (group: F_1 < 1; F_2 (1, 192) = 3.65, MSe = 6895.72, p = .06). Crucially, however, the **priming type*phase*group interaction** seen above was retained **(F_1 (1, 8) = 7.02, MSe = 619.98, p < .05; F_2 (1, 192) = 7.14, MSe = 7047.51, p < .01**; also, a non-significant priming type*phase interaction: F_1 (1, 8) = 3.55, MSe = 619.98, p = .1; F_2 < 1; rest of Fs, F < 1).

2 groups (Crus I vs no TMS): two-way ANOVAs: Associative Priming (group, phase): The two groups did not differ in associative priming sizes across phases (group: F_1 (1, 8) = 1.53, MSe = 355.74, p > .25; F_2 (1, 96) = 1.95, MSe = 7574.45, p = .17). However, associative priming size increased in the second phase across groups (phase: F_1 (1, 8) = 12.88, MSe = 323.35, p < .01; F_2 (1, 96) = 5.75, MSe = 5438.87, p < .05). Crucially, **the increase was stronger for the Crus I group**, significantly by subjects, and marginally by items **(phase*group: F_1 (1, 8) = 6.34, MSe = 323.35, p < .05; F_2 (1, 96) = 3.59, MSe = 7613.36, p = .06).**

2 groups (Crus I vs no TMS): two-way ANOVAs: Categorical Priming (group, phase): Categorical priming sizes did not differ across groups or phases (group: both ps, p > .2; phase: F_1 < 1; F_2 (1, 96) = 1.80, MSe = 6540.00, p = .18). There was, however, a marginal by-items phase*group interaction (F_1 (1, 8) = 2.72, MSe = 849.44, p = .14; F_2 (1, 96) = 3.55, MSe = 6481.65, p = .06), due to the large categorical priming size of the Crus I group in baseline conditions.

2 groups (Crus I vs no TMS): two-way ANOVAs: First phase (group, priming type): A comparison of priming sizes in the first phase showed a marginal group*priming type interaction (F_1 (1, 8) = 4.47, MSe = 540.91, p = .07; F_2 (1, 192) = 3.16, MSe = 6241.01, p = .08), due to the high value of categorical priming observed in

the Crus I group. Priming sizes did not differ between the two groups (group: F_1 (1, 8) = 1.57, MSe = 472.88, $p > .2$; F_2 (1, 192) = 1.78, MSe = 6241.01, $p = .18$).

2 groups (Crus I vs no TMS): two-way ANOVAs: Second phase (group, priming type): Crucially, a comparison of priming sizes in the second half showed a **significant by-items group*priming type interaction (F_1 (1, 8) = 3.14, MSe = 620.40, $p = .11$; F_2 (1, 192) = 4.05, MSe = 7702.22, $p < .05$)**. Priming sizes did not differ between the two groups in the second phase (group: F_1 (1, 8) = 1.3, MSe = 483.86, $p > .25$; F_2 (1, 192) = 1.84, MSe = 7702.22, $p = .18$).

2 groups (Crus I vs no TMS): one-way ANOVAs (group): Comparing priming sizes in each condition, the two groups did not differ in **associative priming sizes** in the **first phase** (group: both *F*s, $F < 1$), but **associative priming size was significantly larger for the Crus I group in the second phase (group: F_1 (1, 8) = 15.7, MSe = 149.78, $p < .005$; F_2 (1, 96) = 5.29, MSe = 8259.19, $p < .05$)**. **Categorical priming size** in the **first phase** was, unexpectedly, significantly larger for the Crus I group (F_1 (1, 8) = 6.02, MSe = 484.48, $p < .05$; F_2 (1, 96) = 5.44, MSe = 5553.39, $p < .05$), while no such difference was observed in the **second phase** (both *F*s, $F < 1$).

The analysis now turns to a comparison of priming sizes **between the Crus I and NCBV TMS groups**.

2 groups (Crus I vs NCBV): A **three-way ANOVA (group, priming type, phase)** showed smaller overall priming sizes for the NCBV group across phases (group: F_1 (1, 8) = 5.79, MSe = 383.04, $p < .05$; F_2 (1, 192) = 5.26, MSe = 8410.09, $p < .05$). **Crucially, the selective increase in associative priming size in the second phase for the Crus I group remained here significant by subjects (priming type*phase*group: F_1 (1, 8) = 5.90, MSe = 761.92, $p < .05$; F_2 (1, 192) = 2.71, MSe = 7216.98, $p = .1$)**. **There was also a significant by-items phase*group interaction (F_1 (1, 8) = 2.04, MSe = 801.96, $p = .19$; F_2 (1, 192) = 5.59, MSe = 7216.98, $p < .05$), due**

to the lack of increase in priming sizes in the NCBV group in the second phase. There was no priming type*phase interaction across groups (F_1 (1, 8) = 2.75, MSe = 761.92, p = .14; F_2 (1, 192) = 2.11, MSe = 8369.77, p = .15; rest of Fs, F < 1, or p > .2).

2 groups (Crus I vs NCBV): two-way ANOVAs: Associative Priming (group, phase): Associative priming size increased only selectively for the Crus I group (phase*group: F_1 (1, 8) = 11.74, MSe = 492.44, p < .01; F_2 (1, 96) = 7.68, MSe = 7552.60, p < .01; phase: F_1 (1, 8) = 2.32, MSe = 492.44, p = .17; F_2 (1, 96) = 1.4, MSe = 7344.51, p > .2; group: F_1 (1, 8) = 2.86, MSe = 137.73, p = .13; F_2 = 2.32, p = .13).

2 groups (Crus I vs NCBV): two-way ANOVAs: Categorical Priming (group, phase): On the other hand, no increase was observed for categorical priming (phase, group*phase: all Fs, F < 1). The Crus I group only showed marginally stronger categorical priming across phases (group: F_1 (1, 8) = 4.38, MSe = 498.82, p = .07; F_2 (1, 96) = 2.97, MSe = 8020.15, p = .09).

2 groups (Crus I vs NCBV): two-way ANOVAs: first half (group, priming type): A comparison for the phase before TMS showed a significant by-subjects group*priming type interaction (F_1 (1, 8) = 6.41, MSe = 577.86, p < .05; F_2 (1, 192) = 1.3, MSe = 7785.52, p > .2; group: both Fs, F < 1). This was due to the poorly matched baseline conditions across groups, and, in particular, to the larger associative priming sizes for the NCBV group in the first phase (see below).

2 groups (Crus I vs NCBV): two-way ANOVAs: second half (group, priming type): For the phase after TMS, **the NCBV group showed smaller sizes, significantly by items, and marginally by subjects (group: F_1 (1, 8) = 4.33, MSe = 885.34, p = .07; F_2 (1, 192) = 10.78, MSe = 7841.55, p < .005)** independently of priming type (group*priming type: F_1 (1, 8) = 2.64, MSe = 437.56, p = .14; F_2 (1, 192) = 1.2, MSe = 7841.55, p > .25).

2 groups (Crus I vs NCBV): one-way ANOVAs (group): Comparing priming sizes of the two groups for each condition, the NCBV group showed larger **associative**

priming size **before TMS** than the Crus I group, but only significantly so by subjects (F_1 (1, 8) = 6.38, MSe = 247.21, $p < .05$; $F_2 < 1$). Crucially, though, **the Crus I group involved larger associative priming size in the after TMS in a much stronger fashion (F_1 (1, 8) = 12.00, MSe = 382.97, $p < .01$; F_2 (1, 96) = 9.34, MSe = 7888.40, $p < .005$).** For **categorical priming**, there was only a weak tendency by items for larger sizes for the Crus I group (**first phase**: F_1 (1, 12) = 3.41, MSe = 630.31, $p = .1$; $F_2 < 1$; **second phase**: $F_1 < 1$; F_2 (1, 96) = 2.50, MSe = 7794.71, $p = .12$).

The analysis now turns to a comparison of priming sizes **between the no TMS and the NCBV TMS groups.**

2 groups (NCBV vs no TMS): three-way ANOVA (priming type, phase, group): A comparison of the priming sizes between the two groups showed a significant by-items **phase*group interaction (F_1 (1, 8) = 2.93, MSe = 602.19, $p = .11$; F_2 (1, 192) = 6.68, MSe = 5474.56, $p < .05$)**, independently of priming type (rest of *p*s, $p > .25$).

2 groups (NCBV vs no TMS): two-way ANOVAs: first phase (priming type, group): The two groups did not differ in priming sizes for the first half of the session (all *p*s, $p > .2$).

2 groups (NCBV vs no TMS): two-way ANOVAs: second phase (priming type, group): In the second half, however, **the NCBV group involved significantly (by items) smaller sizes (group: F_1 (1, 8) = 2.2, MSe = 634.08, $p = .18$; F_2 (1, 192) = 5.70, MSe = 5172.67, $p < .05$)** independently of priming type (group*priming type: both *F*s, $F < 1$, or $p > .2$).

The analysis now turns to the ANOVAs performed **for each of the three groups separately.**

Crus I group: A **two-way ANOVA (priming type, phase)** for the Crus I group alone showed only a **selective increase of associative priming sizes**, marginally by

subjects[98] and significantly by items **(priming type*phase: F_1 (1, 4) = 5.90, MSe = 1078.69, p = .07; F_2 (1, 192) = 4.64, MSe = 8011.52, p < .05**; phase: F_1 (1, 4) = 1.20, MSe = 752.58, p > .3; F_2 (1, 192) = 2.79, MSe = 8011.52, p = .1; rest of ps, p > .25).

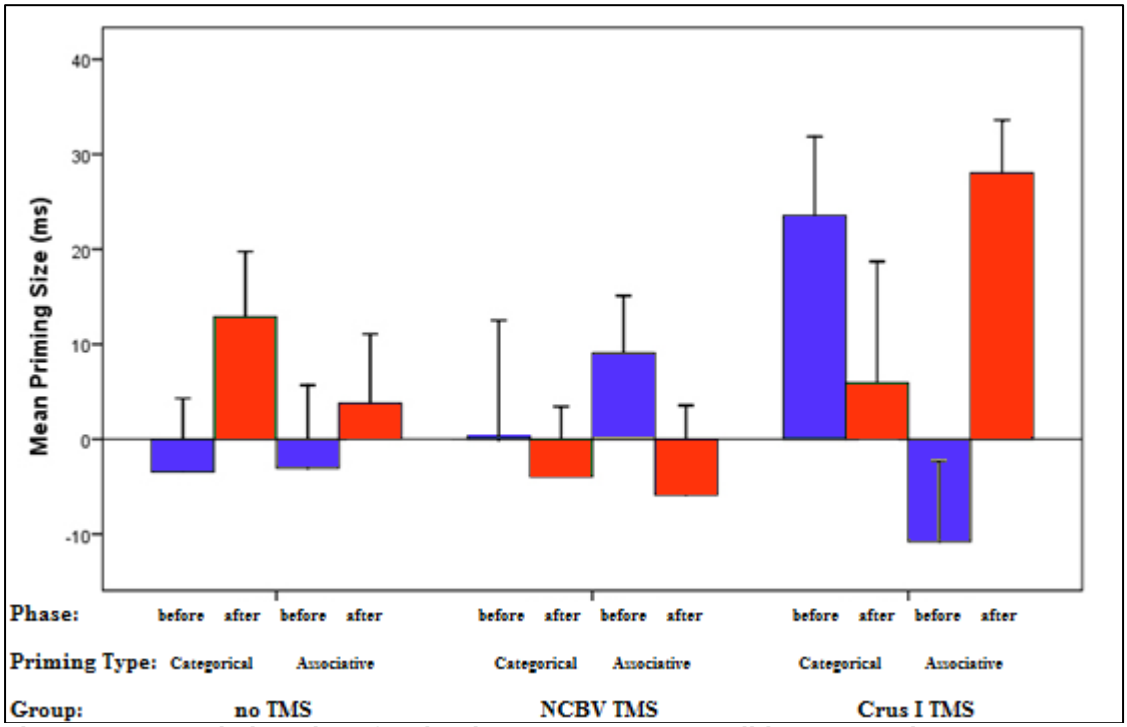
Crus I group: one-way ANOVAs (phase): The interaction above was due to the strong boost of **associative priming** in the second phase (**phase: F_1 (1, 4) = 21.36, MSe = 282.24, p < .05; F_2 (1, 96) = 7.79, MSe = 7517.37, p < .01**), not shown for **categorical priming** (phase: both Fs, F < 1).

No TMS group: A **two-way ANOVA (priming type, phase)** for the no TMS group alone showed a marginal by-items increase in priming sizes in the second phase (phase: F_1 (1, 4) = 3.04, MSe = 353.04, p = .16; F_2 (1, 192) = 3.67, MSe = 5025.42, p = .06), which was marginally stronger by items for categorical priming (priming type*phase: F_1 < 1; F_2 (1, 192) = 3.08, MSe = 5025.42, p = .08).

No TMS group: one-way ANOVAs (phase): An analysis for **categorical priming** demonstrated a marginal by-subjects and significant by-items increase of size in the second phase (phase: F_1 (1, 4) = 7.14, MSe = 149.85, p = .06; F_2 (1, 96) = 7.50, MSe = 4515.98, p < .01), which was not found **for associative priming** sizes (phase: both Fs, F < 1).

NCBV group: two-way ANOVA (priming type, phase): Unlike the other two groups, the NCBV group alone involved **no increase in priming size in the second phase**; on the contrary, it showed **a nonsignificant decrease by items (phase: F_1 < 1; F_2 (1, 192) = 2.39, MSe = 7575.23, p = .12)**, independently of priming type (priming type*phase: both Fs, F < 1, or p > .3).

[98] This interaction, along with a number of other effects, was significant without including the dummy variable.



Plot 7.3: Mean priming sizes for the three groups per condition. Errors bars represent + 1 SEM (experiment 7).

Priming	TMS Phase	Group	Mean Priming Sizes (ms)	SD (ms)
Associative	Before TMS (first half)	No TMS	-3.03	24.64
		NCBV TMS	9.08	17.02
		Crus I TMS	-10.78	18.71
	After TMS (second half)	No TMS	3.79	20.61
		NCBV TMS	-5.85	26.60
		Crus I TMS	28.04	15.73
Categorical	Before TMS (first half)	No TMS	-3.47	21.97
		NCBV TMS	0.34	34.46
		Crus I TMS	23.53	23.52
	After TMS (second half)	No TMS	12.88	19.48
		NCBV TMS	-3.95	20.89
		Crus I TMS	5.93	36.19

Table 7.5: Mean priming sizes for the three groups per condition (experiment 7).

7.4.6.1.1.2. Comparison with the Pilot Group

Since the pilot group (n = 40) used in the previous experiment (section 7.3) completed the experimental session virtually in the same way that the 'no TMS' group (n = 8) did, it was possible to compare this group with the two TMS groups. Mean priming sizes for the new comparison are shown below in table 7.6 and plot 7.4. As above, the analysis will start with the comparisons among the three groups (from three- down to one-way ANOVAs), and terminate with the comparisons between groups. The comparison between the NCBV and the Crus I group, along with analyses for each separate group may be found above.

The comparison of priming sizes among the NCBV (n = 8), Crus I (n = 8) and pilot groups (n = 40) demonstrated effect sizes and significance levels close to those above, especially in analyses by subjects.[99]

3 groups: A three-way ANOVA (priming type, phase, group) showed a significant by-items main effect of group, with the NCBV group having the lowest overall sizes, and the Crus I group the largest ones (group: F_1 (2, 44) = 2.01, MSe = 737.48, p = .15; F_2 (2, 384) = 4.64, MSe = 5925.88, p < .05). The analysis also demonstrated a marginal by-subjects priming type*phase interaction (F_1 (2, 44) = 3.39, MSe = 687.43, p = .07; F_2 (1, 192) = 2.23, MSe = 6606.27, p = .14). Most importantly, however, it showed a significant by-items **phase*group interaction (F_1 (2, 44) = 2.20, MSe = 505.68, p = .12; F_2 (2, 384) = 4.20, MSe = 5574.40, p < .05)**, with sizes not increasing for the NCBV group in the second phase, unlike the other two groups. The analysis also showed a significant by-subjects **priming type*phase*group interaction (F_1 (2, 44) = 3.95, MSe = 687.43, p < .05; F_2 (2, 384) = 2.06, MSe = 5574.40, p = .13**; rest of Fs, F < 1), which was at least partly due to the selective increase of semantic-associative priming after Crus I TMS.

[99] An apparent disadvantage of this comparison is that the variances for the pilot group (n = 40), would be much better-stabilized than those for each TMS group (n = 8). This is the reason why the comparison among the three smaller (n = 8) groups was reported separately.

3 groups: Two-way ANOVAs: Categorical Priming (phase, group): An analysis of categorical priming sizes in particular demonstrated an only marginal by-items difference among the three groups independently of phase, with the Crus I group involving the highest values and the NCBV group the lowest ones (group: F_1 (2, 44) = 1.75, MSe = 705.97, p = .19; F_2 (2, 384) = 2.66, MSe = 5502.29, p = .07). No changes in categorical priming size were found in the second phase (phase, phase*group: all Fs, F < 1, or p > .2).

3 groups: Two-way ANOVAs: Associative Priming (phase, group): On the contrary, the three groups differed significantly in the changes observed in associative priming sizes **(phase*group: F_1 (2, 44) = 5.28, MSe = 548.98, p < .01; F_2 (2, 192) = 5.07, MSe = 5744.39, p < .01)**, with the NCBV group involving no increase, and the Crus I group demonstrating the strongest one in the second phase. The tendency for increase in associative priming across groups in the second phase was only marginal (F_1 (2, 44) = 3.82, MSe = 548.98, p = .06; F_2 (1, 96) = 3.39, MSe = 5400.54, p = .07).The three groups did not differ significantly in associative priming size across phases (group: F_1 < 1; F_2 (2, 192) = 2.03, MSe = 6349.48, p = .13).

3 groups: Two-way ANOVAs: first phase (priming type, group): Comparing priming sizes in the first half showed no significant difference among groups (group: both ps, p > .2); there was only a marginal-by subjects group*priming type interaction (F_1 (2, 44) = 2.64, MSe = 773.14, p = .08; F_2 (2, 384) = 1.2, MSe = 5620.05, p > .3), due to the large categorical priming sizes of the Crus I group in the first phase.

3 groups: Two-way ANOVAs: second phase (priming type, group): A comparison of priming sizes in the second half demonstrated that **the NCBV involved the lowest sizes of the three groups**, marginally by subjects and significantly by items **(group: F_1 (2, 44) = 3.04, MSe = 630.54, p = .06; F_2 (2, 384) = 7.21, MSe = 5880.24, p < .005).** However, as in the analysis of the previous section**, the group*priming type**

interaction in the comparison for the second phase remained elusive (F_1 (2, 44) = 1.8, MSe = 481.67, p = .18; F_2 < 1).

3 groups: One-way ANOVAs: Categorical Priming (group): Comparing categorical priming size per phase, the groups differed only marginally in the **first half** (group: F_1 (2, 44) = 2.71, MSe = 681.04, p = .08; F_2 (2, 192) = 2.53, MSe = 669.06, p = .09), with the Crus I group showing the highest values; no differences were found in the **second half** (both *F*s, F < 1, or p > .2).

3 groups: One-way ANOVAs: Associative Priming (group): On the contrary, examining associative priming sizes per phase, **the Crus I group showed the largest size only in the second half (F_1 (2, 44) = 5.57, MSe = 443.16, p < .01; F_2 (2, 192) = 6.21, MSe = 6066.96, p < .005)**; no differences were found in the **first half** (both *p*s, p > .3).

The analysis now turns to a comparison of priming sizes **between the Crus I and pilot groups.**

2 groups (Crus I vs pilot): A **three-way ANOVA (priming type, phase, group)** showed an increase in priming size in the second half across priming types (phase: F_1 (1, 40) = 4.72, MSe = 471.12, p < .05; F_2 (1, 192) = 6.39, MSe = 5129.11, p < .05), independently of group (phase*group: both *F*s, F < 1). It also demonstrated larger overall priming sizes for the Crus I group, marginally by subjects and significantly by items (group: F_1 (1, 40) = 3.41, MSe = 768.32, p = .07; F_2 (1, 192) = 6.81, MSe = 5590.35, p < .05). This difference was not selective for a priming type (priming type*group: both *F*s, F < 1). Importantly, though, a strong increase in priming size was noted, selectively for associative priming across groups (priming type*phase: F_1 (1, 40) = 8.79, MSe = 711.66, p < .01; F_2 (1, 192) = 4.50, MSe = 5129.11, p < .05; see below), which was **stronger for the Crus I TMS group, significantly by subjects, and marginally by items (priming type*phase*group: F_1 (1, 40) = 6.23, MSe = 711.66, p < .05; F_2 (1, 192) = 2.88, MSe = 5050.74, p = .09).**

2 groups (Crus I vs pilot): Two-way ANOVAs: Associative Priming (phase, group): The two groups did not differ significantly in associative priming size across phases (group: F_1 (1, 40) = 1.2, MSe = 639.61, $p > .25$; F_2 (1, 96) = 2.83, MSe = 6439.85, $p = .1$). There was, however, a strong increase in the second phase across groups (phase: F_1 (1, 40) = 14.93, MSe = 533.62, $p < .001$; F_2 (1, 96) = 11.84, MSe = 4681.17, $p < .005$), which was **significantly stronger by subjects for the Crus I group (phase*group: F_1 (1, 40) = 5.17, MSe = 533.62, $p < .05$; F_2 (1, 96) = 2.36, MSe = 4828.39, $p = .13$).**

2 groups (Crus I vs pilot): Two-way ANOVAs: Categorical Priming (phase, group): The Crus I group unexpectedly showed significantly (by items) larger categorical priming sizes overall (group: F_1 (1, 40) = 2.59, MSe = 714.08, $p = .12$; F_2 (1, 96) = 4.19, MSe = 4740.85, $p < .05$), but not selectively for a phase (phase*group: F_1 (1, 40) = 2.68, MSe = 649.16, $p = .11$; $F_2 < 1$). No change in categorical priming size was observed independently of group (phase: both Fs, $F < 1$).

2 groups (Crus I vs pilot): Two-way ANOVAs: First half (priming type, group): The priming sizes of the two groups in the first phase did not differ across priming types (group: F_1 (1, 40) = 1.46, MSe = 649.41, $p > .2$; F_2 (1, 192) = 2.73, MSe = 4982.45, $p = .1$), but only in a marginally by subjects selective fashion for categorical priming (group*priming type: F_1 (1, 40) = 3.70, MSe = 785.43, $p = .06$; F_2 (1, 192) = 1.9, MSe = 4982.45, $p > .2$), with the Crus I group showing the largest sizes (see below).

2 groups (Crus I vs pilot): Two-way ANOVAs: Second half (priming type, group): Comparing the priming sizes in the second phase between the two groups showed that the Crus I group involved significantly larger priming sizes by items across priming types (group: F_1 (1, 40) = 2.93, MSe = 590.04, $p = .1$; F_2 (1, 192) = 4.49, MSe = 5658.64, $p < .05$). This difference was **larger for the associative priming type, but only marginally by subjects (priming type*group: F_1 (1, 40) = 3.17, MSe = 511.58, $p = .08$; F_2 (1, 192) = 1.4, MSe = 5658.64, $p > .25$).**

2 groups (Crus I vs pilot): One-way ANOVAs: Associative Priming (group): A comparison of the two groups in associative priming sizes per phase showed larger priming size for the Crus I group in the **second half** (group: F_1 (1, 40) = 7.98, MSe = 419.85, $p < .01$; F_2 (1, 96) = 5.00, MSe = 5841.67, $p < .05$), and not for the **first half** (group: both *F*s, $F < 1$).

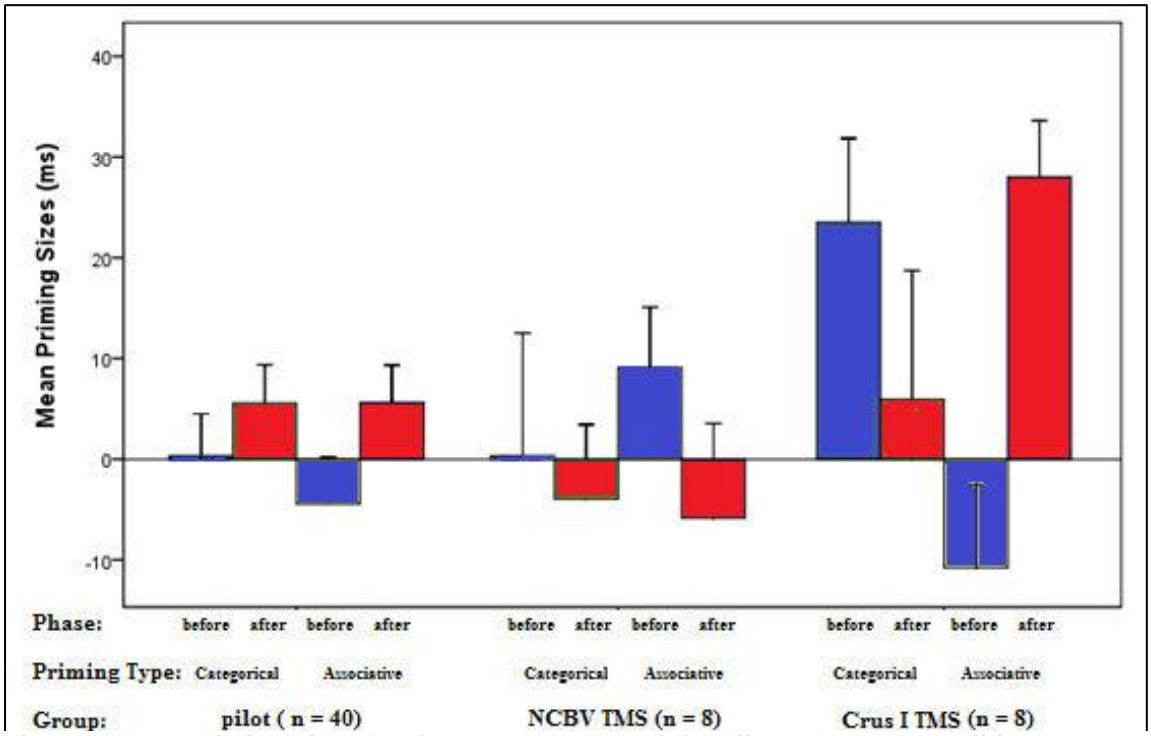
2 groups (Crus I vs pilot): One-way ANOVAs: Categorical Priming (group): The marginally significant differences between the two groups in categorical priming sizes were due to the unexpectedly larger sizes for the Crus I group in the **first**, baseline phase (group: F_1 (1, 40) = 5.27, MSe = 681.47, $p < .05$; F_2 (1, 96) = 4.62, MSe = 4538.34, $p < .05$), but not for the **second** one (group: both *F*s, $F < 1$).

The analysis now turns to a comparison of priming sizes **between the NCBV and pilot groups.**

2 groups (NCBV vs pilot): A three-way ANOVA (priming type, phase, group) showed **no increase in priming sizes after NCBV TMS (phase*group: F_1 (1, 40) = 4.13, MSe = 481.00, $p < .05$; F_2 (1, 192) = 6.52, MSe = 4455.49, $p < .05$)**, no matter the priming type (rest of *F*s, $F < 1$).

2 groups (NCBV vs pilot): Two-way ANOVAs: second half (priming type, group): The interaction above was due to the **significantly (by items) smaller priming sizes in the second half for the NCBV TMS group (group: F_1 (1, 40) = 2.37, MSe = 620.08, $p = .13$; F_2 (1, 192) = 4.16, MSe = 4140.53, $p < .05$)**, once again across priming types (group*priming type: *F*s, $F < 1$).

2 groups (NCBV vs pilot): Two-way ANOVAs: first half (priming type, group): These differences were only marginal by items in the first half of the session (group: $F_1 < 1$; F_2 (1, 192) = 2.95, MSe = 4092.17, $p = .09$), and were not selective for a priming type (group*priming type: both *F*s, $F < 1$).



Plot 7.4: Mean priming sizes for the TMS groups and the pilot group per condition. Error bars represent + 1 SEM (experiment 7).

Priming Type	Phase	Group	Mean Priming Sizes (ms)	SD (ms)
Categorical	First Half	Pilot	0.33	26.37
		NCBV TMS	0.34	34.46
		Crus I TMS	23.53	23.52
	Second Half	Pilot	5.56	24.21
		NCBV TMS	-3.95	20.89
		Crus I TMS	5.93	36.19
Associative	First Half	Pilot	-4.44	29.40
		NCBV TMS	9.08	17.02
		Crus I TMS	-10.78	18.71
	Second Half	Pilot	5.62	23.43
		NCBV TMS	-5.85	26.60
		Crus I TMS	28.04	15.73

Table 7.6: Mean priming sizes for the TMS groups and the pilot group per condition (experiment 7).

7.4.6.1.1.3. Comparison of Thematic Roles

Given that the associative priming set involved four different θ-role subsets (agent, instrument, location, patient; see sections 7.2.1 and 7.3.1), a number of additional analyses were performed to determine whether the TMS effects observed were selective for a subset. Comparing the changes in the priming sizes of the four different subsets among the NCBV (n = 8), Crus I (n = 8), and No TMS (n = 8) groups, the increase was shown to be selectively stronger for the Crus I group, but independent of subset (phase*group: F_1 (2, 12) = 8.35, MSe = 1550.47, $p < .01$; F_2 (2, 168) = 5.22, MSe = 6918.89, $p < .01$; θ-role*phase*group: both *F*s, $F < 1$). A comparison of different θ-role priming sizes for that group showed an in the second phase independently of subset (phase: F_1 (1, 4) = 32.54, MSe = 768.92, $p < .01$; F_2 (1, 84) = 10.47, MSe = 7413.31, $p < .005$; phase*θ-role: $F_1 < 1$; F_2 (3, 84) = 1.80, MSe = 7413.31, $p = .15$). No such increase in associative priming size was observed in the NCBV or the pilot group (all *F*s, $F < 1$, or $p > .25$).

Similarly, comparing the priming sizes of the different subtypes among the NCBV (n =8), Crus I (n = 8), and pilot groups (n = 40) demonstrated that the increase is stronger for the Crus I group, independently of subset (phase*group: F_1 (2, 44) = 4.84, MSe = 2669.55, $p < .05$; F_2 (2, 168) = 5.99, MSe = 5804.53, $p < .005$; phase*group*θ-role: both *F*s, $F < 1$).

7.4.6.1.2. Lexical Decision Accuracy

As in the previous studies, priming-related effects on accuracy were calculated as differences between the arcsine-transformed accuracy rates for unrelated items and those for related items per group (section 4.4.11). A three-way ANOVA over these differences (group, priming type, phase) showed an only marginal by-subjects increase in the difference between unrelated and related items in the second phase (phase: F_1 (2, 12) =

3.81, MSe = 0.06, p = .08; F_2 (1, 192) = 2.62, MSe = 0.22, p = .11), independently of priming type and/ or group (rest of Fs, F < 1, or p > .2).

7.4.6.2. Effects on Sensorimotor Performance

7.4.6.2.1. Lexical Decision Latencies

An analysis of lexical decision latencies for words and nonwords for the three groups (n = 8 each) in the first and second half of the session (plot 7.5 and table 7.7 below) showed a much stronger decrease in RTs for the 'no TMS' group in the second phase than for the TMS ones. As above, the analysis will begin with the ANOVAs involving all three groups (three-, and, wherever necessary, two- and one-way), and will proceed to comparisons between two groups, terminating with the appropriate ANOVAs for separate groups.

3 groups: A three-way ANOVA (group, wordness, phase) demonstrated a clear difference between words and nonwords across groups, with words being judged as such much more quickly than nonwords (wordness: F_1 (1, 12) = 24.85, MSe = 967.90, p < .001; F_2 (1, 492) = 130.56, MSe = 5961.83, p < .001). Unexpectedly, latencies differed significantly (by items) among groups (group: F_1 < 1; F_2 (2, 984) = 32.02, MSe = 1984.27, p < .001), with the Crus I group involving the longest ones. Latencies also became shorter in the second phase across groups (phase: F_1 (1, 12) = 27.74, MSe = 494.49, p < .001; F_2 (1, 492) = 169.34, MSe = 2262.25, p < .001). The difference between words and nonwords is stronger in the no TMS group, significantly so by items (wordness*group: F_1 (2, 12) = 1.4, MSe = 967.90, p > .3; F_2 (2, 984) = 30.44, MSe = 1984.83, p < .001). The difference in latencies between words and nonwords changes significantly (by items) in the second phase, with words receiving even shorter latencies in the second phase than nonwords (wordness*phase: F_1 < 1; F_2 (1, 492) = 7.28, MSe = 2262.25, p < .01). However, crucially for the analysis here, **latencies for the two TMS groups became shorter in the second phase to a much smaller extent than those in**

the no TMS group, across target types. This is demonstrated by a significant by-items phase*group interaction (F_1 (2, 12) = 1.7, MSe = 494.49, p > .2; F_2 (2, 984) = 12.73, MSe = 2023.47, p < .001; wordness*phase*group: both *F*s, $F < 1$ or $p > .3$).[100]

3 groups: Two-way ANOVAs: First phase (group, wordness): Given that the analysis was between subjects with a small size for each subject group (n = 8), latencies were not well-matched in the baseline conditions across groups: the three groups differed significantly (by items) in the first phase (group: $F_1 < 1$; F_2 (2, 984) = 17.02, MSe = 2162.94, $p < .001$), with the Crus I group showing the longest, and the NCBV group the shortest ones. The analysis also showed a main effect of wordness, as above (F_1 (1, 12) = 20.92, MSe = 555.44, $p < .005$; F_2 (1, 492) = 116.25, MSe = 4392.28, $p < .001$), and a wordness*group interaction ($F_1 < 1$; F_2 (2, 984) = 12.29, MSe = 2162.94, $p < .001$), observed also in the second phase, with the no TMS group showing the greatest differences in latencies between words and nonwords.

3 groups: Two-way ANOVAs: Second phase (group, wordness): The groups **differed significantly (by items) in the second phase with larger effect sizes (group: $F_1 < 1$; F_2 (2, 984) = 28.45, MSe = 1844.79, $p < .001$)**, with the no TMS group showing the shortest latencies. The analysis also showed the same main effect of wordness (F_1 (1, 12) = 18.01, MSe = 687.22, $p < .005$; F_2 (1, 492) = 74.17, MSe = 3831.81, $p < .001$), and the same wordness*group interaction (F_1 (2, 12) = 2.06, MSe = 687.22, $p = .17$; F_2 (2, 984) = 18.50, MSe = 1844.79, $p < .001$), with words receiving shorter latencies than nonwords, and even more so for the no TMS group.

[100] In view of space restrictions, a number of ANOVAs are not reported. As there was no wordness*phase*group interaction, the separate analyses across groups for words and nonwords are not reported here. Both these analyses showed the same group*phase interaction as the one above. Similarly, because of the fact that groups differed significantly overall in latencies, further comparisons between groups separately in each phase are not reported. In all of these comparisons, groups differed both in the first and in the second phase. However, the effect size of the difference between any of the two TMS groups and the no TMS one is larger in the phase after than in the phase before. For all of these cases, the group*phase interaction is reported instead.

The analysis now moves to a comparison of latencies **between the NCBV and no TMS groups.**

2 groups (NCBV vs no TMS): A **three-way ANOVA (group, phase, wordness)** showed once again a significant by-items main effect of group, with the NCBV group involving the shortest latencies of the two (group: $F_1 < 1$; F_2 (1,492) = 10.15, MSe = 2017.53, $p < .005$). As above, words were judged more rapidly than nonwords (wordness: F_1 (1, 8) = 15.96, MSe = 1252.15, $p < .005$; F_2 (1, 492) = 140.00, MSe = 4758.07, $p < .001$), and even more strongly for the no TMS group. This is shown by a significant by-items group*wordness interaction (F_1 (1, 8) = 1.4, MSe = 1252.15, $p > .25$; F_2 (1, 492) = 46.22, MSe = 2017.53, $p < .001$). As above, latencies became much shorter across groups in the second phase (phase: F_1 (1, 8) = 18.06, MSe = 672.53, $p < .005$; F_2 (1, 492) = 140.22, MSe = 2400.94, $p < .001$), and even more so for word targets- this is demonstrated by a significant by-items phase*wordness interaction ($F_1 < 1$; F_2 (1, 492) = 4.55, MSe = 2400.94, $p < .05$). Crucially, though, there was **a significant by-items group*phase interaction (F_1 (1, 8) = 1.3, MSe = 672.53, $p > .25$; F_2 (1, 492) = 16.54, MSe = 2099.07, $p < .001$), with the NCBV group showing a much smaller decrease of RTs in the second phase**, independently of target type (wordness, and rest of *p*s, $p > .2$).

The analysis now moves to the comparison of latencies **between the no TMS and the Crus I group.**

2 groups (Crus I vs no TMS): A **three-way ANOVA (group, phase, wordness)** showed that words were also judged more rapidly than nonwords (wordness: F_1 (1, 8) = 21.63, MSe = 904.72, $p < .005$; F_2 (1, 492) = 139.29, MSe = 4888.01, $p < .001$), and even more strongly for the no TMS group- this was demonstrated by a significant by-items group*wordness interaction (F_1 (1, 8) = 2.02, MSe = 904.72, $p = .19$; F_2 (1, 492) = 44.72, MSe = 1964.62, $p < .001$). As demonstrated by a significant by-items phase*wordness interaction ($F_1 < 1$; F_2 (1,492) = 3.95, MSe = 2169.37, $p < .05$),

this wordness difference became larger in the second phase, independently of group (no group*phase*wordness interaction: F_1 (1, 8) = 3.04, MSe = 153.07, p = .12; F_2 < 1).As above, the no TMS group showed shorter latencies overall (group: F_1 < 1; F_2 (1, 492) = 22.70, MSe = 1964.62, p < .001). Crucially, however, there was a clear reduction in latencies in the second phase across groups (phase: F_1 (1, 8) = 21.56, MSe = 498.84, p < .005; F_2 (1, 492) = 145.12, MSe = 2169.37, p < .001), which was **significantly smaller for the Crus I group in the second phase- this is demonstrated by a significant by-items group*phase interaction (F_1 (1, 8) = 2.78, MSe = 498.84, p = .13; F_2 (1, 492) = 22.72, MSe = 1857.33, p < .001).**

The analysis now moves to the comparison of latencies **between the NCBV and the Crus I group.**

2 groups (Crus I vs NCBV): Three-way ANOVA (group, phase, wordness): However, a comparison between the latencies of the Crus I and those of the NCBV group showed that **the two groups did not differ on the extent of such decrease of RTs in the second phase (group*phase: both *F*s, *F* < 1)**. The analysis only showed a decrease in RTs across the two groups (phase: F_1 (1, 8) = 17.06, MSe = 312.09, p < .005; F_2 (1, 492) = 71.03, MSe = 1977.65, p < .001). The two groups once again differed by items in overall latencies, with the NCBV group showing the shortest latencies (group: F_1 < 1; F_2 (1, 492) = 63.70, MSe = 1970.65, p < .001). There was also a clear difference in latencies between words and nonwords, with words being judged faster than nonwords (wordness: F_1 (1, 8) = 13.03, MSe =746.83, p < .01; F_2 (1, 492) = 63.39, MSe = 4261.86, p < .001). There was also a phase*wordness interaction by items, with the decrease in RTs for words being larger than that for nonwords (F_1 < 1; F_2 (1, 492) = 6.94, MSe = 1977.65, p < .01; rest of *F*s, *F* < 1).

The analysis now moves to a number of two-way ANOVAs for **each group separately.**

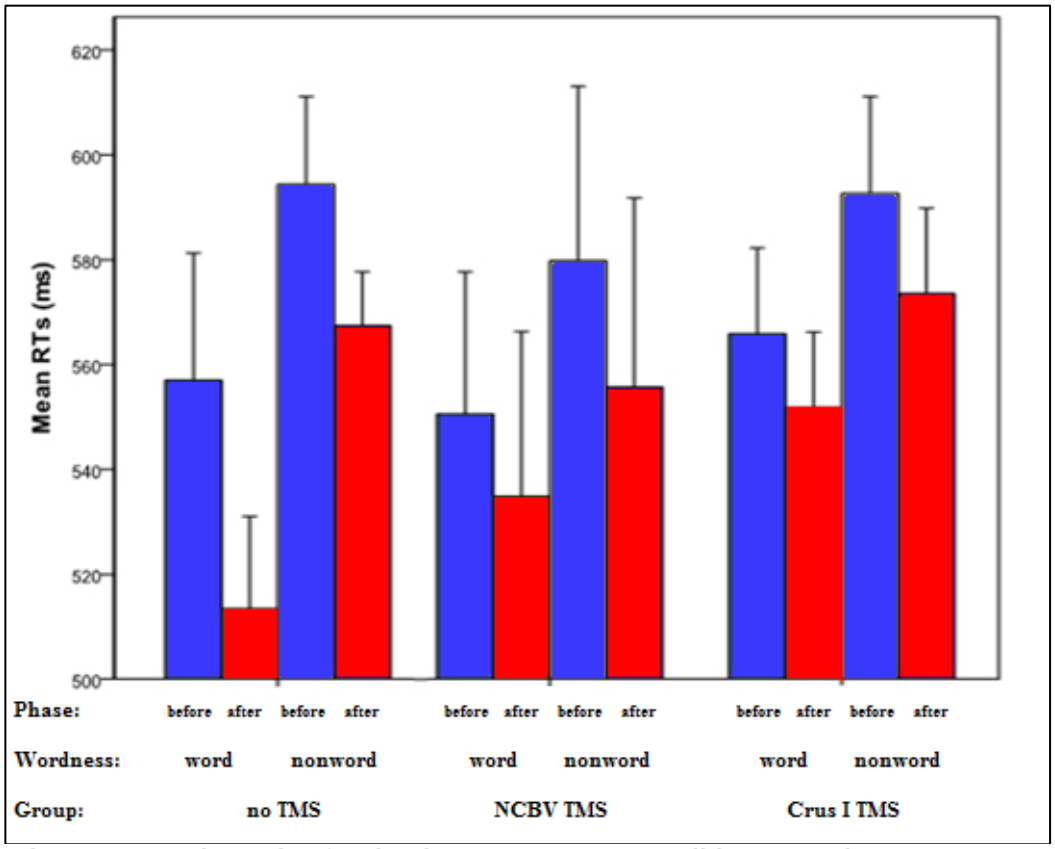
1 group: no TMS: Two-way ANOVAs (phase, wordness): An analysis of latencies for the no TMS group demonstrated a decrease in RTs that is much stronger than that shown for the two TMS groups below (phase: F_1 (1, 4) = 11.56, MSe = 859.28, $p < .05$; F_2 (1, 492) = 132.50, MSe = 2217.53, $p < .001$). This difference in effect sizes grounds the interactions of interest observed above. There was also a main effect of wordness (F_1 (1, 4) = 11.82, MSe = 1410.04, $p < .05$; F_2 (1, 492) = 170.08, MSe = 3697.86, $p < .001$), and a non-significant phase*wordness interaction (F_1 (1, 4) = 2.88, MSe = 190.95, $p = .17$; F_2 (1, 492) = 1.44, MSe = 2217.53, $p > .2$), also observed in marginal levels for the two TMS groups below.

1 group: Crus I: Two-way ANOVA (phase, wordness): An analysis of latencies for the Crus I group showed **a decrease in RTs with effect sizes smaller than those for the no TMS group above (phase: F_1 (1, 4) = 15.96, MSe = 138.41, $p < .05$; F_2 (1, 492) = 34.96, MSe = 1809.17, $p < .001$)**. As above, there was a preserved main effect of wordness (F_1 (1, 4) = 11.82, MSe = 399.41, $p < .05$; F_2 (1, 492) = 44.30, MSe = 3154.77, $p < .001$), and a marginal by-items phase*wordness interaction ($F_1 < 1$; F_2 (1, 492) = 3.06, MSe = 1809.17, $p = .08$).

1 group: NCBV: Two-way ANOVA (phase, wordness): The decrease in RTs for the NCBV group was marginal by subjects and significant by items, with **effect sizes smaller than those for the no TMS group (phase: F_1 (1, 4) = 6.50, MSe = 485.78, $p = .06$; F_2 (1, 492) = 33.99, MSe = 2282.48, $p < .001$)**. As above, there was a marginal by-items wordness effect (F_1 (1, 4) = 4.58, MSe = 1094.26, $p = .1$; F_2 (1, 492) = 42.39, MSe = 3077.74, $p < .001$), and a marginal by-items phase*wordness interaction ($F_1 < 1$; F_2 (1, 492) = 3.65, MSe = 2282.48, $p = .06$).

Target Letter String	Phase (Half)	Group	Mean RTs (ms)	SD (ms)
Word	**First**	**NCBV TMS**	550.58	76.61
		No TMS	557.01	68.62
		Crus I TMS	565.88	46.25
	Second	**NCBV TMS**	534.89	88.75
		No TMS	513.49	49.64
		Crus I TMS	551.76	40.75
Nonword	**First**	**NCBV TMS**	579.79	94.12
		No TMS	594.37	47.49
		Crus I TMS	592.68	52.22
	Second	**NCBV TMS**	555.74	101.79
		No TMS	567.42	29.13
		Crus I TMS	573.56	46.05

Table 7.7: Mean latencies for the three groups per condition (experiment 7).



Plot 7.5: Mean latencies for the three groups per condition. Error bars represent + 1 SEM (experiment 7).

7.4.6.2.2. Lexical Decision Accuracy

Accuracy rates were arcsine-transformed and subjected to a three-way ANOVA (group, phase, wordness), which only showed a non-significant trend by items for higher accuracy rates in the second phase across groups (phase: $F_1 < 1$; $F_2(1, 492) = 1.81$, MSe = 0.06, $p = .18$). However, no change occurred selectively for a group (phase*group: both *p*s, $p > .25$; rest of *p*s, $p > .3$).

7.4.7. Discussion

The results above provide evidence for a rather selective enhancement of semantic-associative priming after stimulation of the right NCB Crus I. While such increase in priming size occurred in the pilot group as well, it was significantly smaller, and was not selective for a priming type. On the contrary, priming did not occur after NCBV TMS, which can be interpreted as a TMS effect on perceptual or cognitive processes. Latencies after TMS of either site were not as short as those in the second phase for the no TMS group.

7.4.7.1. Sensorimotor Effects

The decrease in RTs for lexical decisions in the second phase of the session was much smaller in the two TMS groups than that in the 'no TMS' group. This finding differs from those of the two previous TMS studies (chapters 5 and 6), where it was selectively NCBV TMS that showed this effect. Of course, the power of the analysis was constrained by the between-subjects design here. One possibility is that of a placebo effect after TMS, with slower lexical decisions being made independently of stimulation site. Another possibility is that the effect induced by TMS of the Crus I could be cognitive in nature, and could involve the disruption of practice-induced facilitation, much like in the case of PLCB patients (section 2.4.2). Thus the much smaller decrease in RTs observed after TMS would be owed to a perceptual, NCBV TMS-induced effect,

and a cognitive, Crus I-induced one. Stronger designs with larger subject pools would be required to explore this possibility.

7.4.7.2. Cognitive Effects

The findings here were in line with the predictions made in the second chapter, where it was hypothesized that disruption of CB cortical activity would amount to disinhibited predictions in semantic processing (section 2.6.2.2), much like in the first study, where NCBV TMS enhanced formal-associative priming.

As in the second study (chapter 6), not all interactions reached significance, and many of them showed sizeable effects either by subjects or by items. This may suggest that not all subjects showed the effect of TMS, plausibly because of the lack of precise neuronavigated coordinates and due to the significant depth of the site. It would also suggest that priming was affected for only a subset of items, and/or that the number of subjects and/or items did not suffice to yield significant effects in all cases.

The fact that priming was only observed in the second phase of the session in the preceding pilot experiment might suggest that the large increase in associative priming size for the Crus I group is also strategic in nature. Instead of a direct enhancement of non-attentional semantic-associative priming then, Crus I TMS might have arguably disrupted access to associative memory. In so doing, it might have prompted heavier involvement of the strategic component, which is arguably at work here. In order to validate this interpretation, tasks showing priming independently of phase should be used instead.

In spite of the marginal differences among groups in semantic-categorical priming sizes in baseline conditions, the effect of Crus I TMS was rather selective for semantic-associative, and not for semantic-categorical priming. This is in line with the findings in Fiez *et al.* (1992) and Gebhart *et al.* (2002), where CB patients showed

retained semantic representations and were able to select the appropriate verb for a given noun from a list, despite their deficits in verb generation (section 2.4.4.5). This suggests that the information stored in those PLCB loci is not one of lexical (verb) entries, but pertains to the rapid application of event knowledge in the linguistic stream (Ferretti *et al.*, 2007).

A question in this study would be whether the mechanism affected is one of covert production. In other words, it might be the case that an ‘‘internal generation’’ component is affected here (Gebhart *et al.*, 2002). From this perspective, such θ-role-based priming would be observed because the NCB internally generates a verb-response to the noun-stimulus. As in the first study (section 5.4.7.2.2), the results here cannot discard this interpretation, and only contrastive stimulation of cortical language production-related loci would provide some insight. However, as discussed above, such explanations restrict the role of the CB to the production modality, and ignore its significance in grounding feedforward control and state estimation par excellence (section 2.5).

Another question would be whether a doubly dissociative mapping is tenable between formal- and semantic-associative processes on the one hand, and medial and lateral NCB circuits on the other. The evidence here does not suffice for such an argument. At first blush, the fact that right NCBV TMS did not induce an enhancement of associative priming might be taken to suggest that such a dissociation is tenable. However, the anatomical co-occurrence of the ‘oculomotor vermis’ with the ‘neocerebellar vermis’ complicates the picture- if anything, this was the take-home lesson of the second TMS study, where it might as well have been the case that right NCBV TMS impaired perception of the prime words altogether. In fact, the third TMS experiment used visual settings similar to the second one (chapter 6): both prime and target words appeared in lower-case black letters at the center of a white screen, and participants were sat in a brightly lit room. Thus, the absence of a priming enhancement after right NCBV TMS in comparison with the performance of the pilot group might

have been caused by perceptual disruptions after right NCBV TMS. On the contrary, the visual settings of the first experiment did not impose such heavy perceptual demands: prime words appeared in lower-case-, while target words appeared in upper-case letters; stimuli appeared in green fonts against a black background, and were perceived in a dimly lit room; thus perceptual disruptions accompanying NCBV TMS might have been alleviated to a considerable extent. Hence, it remains an open question whether NCBV TMS in contexts of lower perceptual demands would induce cognitive effects in the same direction as those of Crus I TMS.

A further issue would be whether the enhancement observed here in θ-role-based priming should rely on the same neurocognitive mechanisms as other semantic-associative types. In other words, would the right posterolateral Crus I be specialized particularly in action-oriented associations, prototypically involving an agent (e.g. 'hunter'), a patient (e.g. 'pray'), an instrument (e.g. 'gun') or a location (e.g. 'forest') on the one hand, and an action verb (e.g. 'hunting') on the other? Or would it be as strongly involved in any association of semantic representations (e.g. 'gun'-'hunter', or 'forest'-'pray')? If there is such a specialization, is it due to semantic properties (representation of action), or to grammatical class (verb)? Further experimentation is required to contrast these different kinds of priming.

Much like in the first TMS study (chapter 5), the findings here may be seen in the light of particular psycholinguistic phenomena, as studied in experimental settings (chapter 2), or as shown in historical linguistic data (chapter 3). Such phenomena would rely on semantic-associative computations that can be applied in sentence processing. In the form of schema-transmission, these computations would be involved in garden paths, semantic and verbal illusions, and θ-role misassignments in processing sentences with non-canonical word order (sections 2.4 and 2.6.1). Beyond schema-transmission, semantic-associative computations of the NCBKFLP would include processes like the pragmatic routines and the metonymic inferential shortcuts discussed in the context of grammaticalization changes (section 3.6.3.2).

7.5. Conclusion

The present study assessed semantic-categorical and semantic-associative noun-to-verb priming before and after TMS of different NCB compartments, and before and after no TMS at all. It was shown that stimulation of the right NCB Crus I enhanced associative priming rather selectively, and to a significantly larger extent than any strategic effect did in the pilot group. The present study thus provided some first TMS evidence for the selective involvement of NCB loci in semantic-associative computations. These findings were discussed in the light of the NCBKFLP hypothesis (chapter 2).

Conclusion of Volume II: General Discussion

The empirical part of the thesis described the first steps taken in the experimental exploration of the hypotheses generated in the theoretical part. After presenting the reasons why TMS experimentation was preferred over clinical or imaging studies, the discussion moved to the particular settings of stimulation used (chapter 4). Three TMS studies on the neocerebellum (NCB) were reported, which are, to the best of my knowledge, the first CB TMS studies on language processing.

Despite a considerable number of limiting conditions (summarized in section 8.2 below), the following patterns were observed. The first study (chapter 5) demonstrated a selective enhancement of formal-associative priming after NCBV TMS. The second study demonstrated a marginally selective decrease in semantic-categorical priming sizes after NCBV TMS, along with a clear disruption of the practice-induced reduction of reaction times after NCBV TMS in processing repeated pairs of letter strings. Finally, the third one showed a quite selective increase in semantic-associative priming sizes after TMS of the Crus I, which was larger than that observed in the no TMS group. No categorical or associative priming was shown after NCBV TMS.

TMS on the NCB cortex may have introduced neural noise into the Purkinje cell layer, the output of which inhibits the deep CB nuclei. This would result in disinhibited CB output to the language-related cortical loci. This output may reflect predictions generated on the basis of co-occurrence relations between different linguistic events – in this case, relations between lexical forms, i.e. ‘formal associates’, or of lexical semantic representations, i.e. ‘semantic associates’. NCBV TMS may have also compromised access to the co-occurrence information on the previously encountered pairs of prime-target letter strings, which may reside in the Purkinje cell layer of the CB cortex. However, the first results of theta-burst stimulation (TBS) of the CB have shown effects in the opposite direction from those of 1 Hz CB rTMS (section 4.4.9). Furthermore, it

remains unclear why NCB cTBS enhanced performance in processing associatively related word pairs, while it disrupted performance in processing newly acquired letter strings that co-occurred ad hoc only in experimental settings. Thus, while further research is required to provide a concrete explanation for the direction of the effects, their selective nature remains quite suggestive.

The results also demonstrated some tendencies for reduced semantic-categorical priming after NCBV TMS. The marginal trend in the first study for semantic-categorical priming to drop selectively after NCBV TMS was partly owed to differences in pre-TMS baseline conditions between the two sites (section 5.4.6.1.1). More clearly, in the second study, NCBV TMS impaired semantic-categorical priming in a marginally selective fashion (section 6.2.7.2). In the third study, neither semantic-associative nor semantic-categorical priming gained significant size in the second half after NCBV TMS, unlike sizes for the other groups (section 7.4.7.2). However, the second and third tasks involved higher perceptual demands; thus, a perceptual disruption induced by the NCBV TMS could have impaired the perception of the prime word altogether. On the contrary, the relatively spared formal-associative priming sizes after NCBV TMS in the second study could be explained on the grounds of a selective cognitive enhancement cancelling out the perceptual disruptions induced (section 6.2.7.2).

Furthermore, the studies here provided some first evidence in the literature for disruptions in reading-related processes after NCBV TMS. These were manifested in lower accuracy rates after NCBV TMS in the first session of participation in the first study (section 5.4.7.1), and in longer latencies after NCBV TMS in the second study (section 6.2.7.1). The design of the third study did not allow to dissociate placebo effects from actual TMS effects on performance after stimulation (section 7.4.6.2.1). Finally, there was a tendency for longer latencies after NCBV TMS selectively for ‘word’ responses in the first experiment. This trend could be accounted for by an additional disruption in the control of right-hand index-finger responses after NCBV TMS. As discussed in the next chapter, a number of conditions would need to be met to ensure accurate stimulation of CB compartments to yield more robust effects.

Chapter 8

Conclusions and Future Directions

> ‘‘The unconscious comes before the conscious. The logic of the historic process comes before the subjective logic of the human beings who participate in the historic process’’ (Luxemburg, 1904 [1961, p. 93]).

8.1. Summary

The present work developed a synthesis of neocerebellar (NCB) computations and psycholinguistic models employing state estimation and feedforward control. In this manner, it provided a first, testable answer to the long-standing, yet poorly addressed question of the way in which cerebellar (CB) computations participate in language processing. In particular, it was argued that the NCB provides a strong candidate for the instantiation of associative computations of psycholinguistic mechanisms. These associative operations of the cortico-cerebellar circuits may bias, override, or bypass the categorical operations of cortico-cortical mechanisms. The NCB can also adaptively control the trade-off between cortico-cortical and cortico-cerebellar involvement in the processed linguistic signals. A number of findings were discussed in the light of this hypothesis, and several mechanisms in sentence processing were shown to rely on computations that are fundamentally instantiated by NCB circuitry. The phenomena in which such mechanisms are often captured also provide the testing grounds for the empirical assessment of the hypothesis here (chapters 1 and 2).

The argument was then taken further, to suggest that NCB computations are fundamental for the dialogical routinization phenomena underlying grammaticalization operations. In this vein, an attempt was made to bring neurolinguistics, psycholinguistics, and historical linguistics in contexts of richer interdisciplinary

discourse. Automatization, the domain-general core of grammaticalization processes, was shown to heavily rely on CB circuitry. The uniform and multimodal computations of the latter enable the exploration of NCB contributions to grammaticalization changes. In particular, it was argued that the intra-generational psycholinguistic changes underlying grammaticalization operations rely on an adaptive shift from categorical towards associative computations in the online processing of constructions in routinized dialogical contexts. To this end, a range of phenomena accompanying grammaticalization were shown to involve such associative mechanisms. According to the argument, the NCB would be fundamental for both the associative component, and for the adaptive trade-off between the cortico-cortical categorical and cortico-cerebellar associative component (chapter 3).

Finally, the first steps have been taken here to explore these hypotheses with TMS, constructing and running the first psycholinguistic experiments for CB stimulation (chapter 4). The first study compared formal-associative with semantic-categorical noun-to-noun priming. A selective increase in associative priming was found after TMS of the right Neocerebellar Vermis (NCBV), along with evidence for a reading-related disruption (chapter 5). The second study, also comparing the two priming sizes in Modern Greek, showed evidence for a rather selective impairment in semantic-categorical priming, along with disruptions in reading-related processes after NCBV TMS. This pattern was discussed as a product of the interaction of associative priming enhancement and disrupted perception of the prime word, or as a cognitive effect on the categorical-associative trade-off. Above all, the study demonstrated a lack of shorter latencies after NCBV TMS, selectively in the second session of participation. This was discussed as a result of disrupted linguistic associative memory access/ expression following NCB TMS (chapter 6). The third study compared semantic-categorical with semantic-associative noun-to-verb priming. Whereas priming reached significance only in the second phase for a pilot group, semantic-associative priming increased much more strongly and rather selectively after stimulation of the right NCB Crus I. This was in line with imaging results demonstrating the involvement of this compartment in cognitive

aspects of verb generation. On the contrary, (semantic) priming overall for the NCBV TMS group showed no increase in priming size in the second half (chapter 7), which could be explained as either a perceptual or as a cognitive effect. Despite the weaknesses in the results, owed, to a large extent, to the limiting conditions in experimentation here (see next section), such selective involvement of NCB circuitry in formal and semantic associations provides evidence for the hypothesis that NCB circuits support state estimation and feedforward control in language processing.

8.2. Limiting Conditions and Future Directions

While the absence of further experimental work in the thesis may be frustrating, it is apparent that the hypothesis at hand required significant interdisciplinary theoretical analysis. NCB language functions had not been approached from an emulative perspective in more than a speculative fashion, lacking psycholinguistic insight (section 2.3.1). Similarly, historical linguists remain, with a few exceptions, far from psycholinguistic modelling, and, with even less exceptions, from neurolinguistic research (sections 3.3.2 and 3.3.3).

8.2.1. On TMS Experimentation

Most importantly though, there was no pre-established CB TMS paradigm to follow for psycholinguistic experimentation, since the studies reported here were the first of their kind (section 4.4). Above all, significant time and resources had to be spent in acquiring the necessary expertise for TMS experimentation. The TMS work, including the number of experiments, was further constrained by a number of technical issues:

The unavailability of a complete and in use TMS apparatus in the University of Edinburgh (R. Shillcock, J. Brooks, personal communication, 2008-2010), could only minimally be compensated for by relying on personal contacts with TMS laboratories in other Institutions, especially the Institute of Cognitive Neuroscience (University College London). Other laboratories were often unavailable (e.g. University of Birmingham: Prof. Chris Miall, personal communication, 16 April 2009; Neuroscience and Aphasia Research Unit, University of Manchester: Dr. Gorana Pobric, personal communication, 29 April 2009). On its own, a visiting status in TMS laboratories imposed heavy constraints on the amount of time that a laboratory room could be booked for, both due

to laboratory availability, and due to available resources for longer stays. Limited time thus made subject pools of satisfactory sizes unattainable. Furthermore, the heavily constraining native-speaker requirement (section 4.4.12), especially for the largely international London student population, meant loosening on other restrictions in participant recruitment. No registered MRIs were set as a prerequisite for participation, the obvious cost being that no neuronavigational software (unavailable in the 'Papanikolaou' hospital lab) could be used to accurately map brain- to scalp-coordinates for the localization of the stimulation area. Moreover, the prominence of the inion, used as a landmark here (section 4.4.8), varied greatly across participants, thus compromising the reliable calculation of the scalp-based coordinates. This was especially reflected in the second and third TMS studies, where interactions often reached significance only by items, partly suggesting that only a subset of participants may have been successfully stimulated (see also section 8.2.3). Thus, stronger resources, or long-term placement in research contexts of organized TMS experimentation, would ensure larger subject pools, allowing for accurate, individuated mapping of brain-to-scalp coordinates.

Another limitation which is quite intrinsic to TMS is the complexity in interpreting results given the causal nature of the methodology involved (section 4.2). This translated into a question as to whether the changes induced should be interpreted as due to a secondary, indirect effect on the language-related cortical loci to which the stimulated CB areas ultimately project, or due to a direct effect on CB-based language processes. For example, the sensorimotor effects induced in Miall and Christensen (2004) after CB TMS were seen as products of either a direct CB influence, or an indirect influence via the CB effect on motor cortical function. In the present context, the nature of the CB computations (section 2.2), the suggestive evidence for CB involvement in such tasks (section 2.4), and the absence of any known associations with cortical pathology (e.g. section 2.4.4.4) suggest that the TMS neural noise affected direct NCB contributions. Stronger evidence could be acquired with the addition of experimental conditions including contrastive stimulation of language-related cortical

loci and CB ones. In turn, more experimental conditions would require a significant number of participants and/or stimuli.

The sensory disruptions following NCBV TMS were also difficult to dissociate from any cognitive aspects in the second and third studies (chapters 6 and 7). It could not be determined whether the impairments in semantic priming after NCBV TMS were partly or even completely owed to disruptions in the perception of the prime word. In future work, the use of a perceptually less demanding task or the addition of a control task could provide insight into the extent to which lack of priming may be attributed to sensory or cognitive effects.

Furthermore, the studies here in fact departed considerably from the majority of offline CB TMS experiments with respect to the protocol used, by employing theta-burst stimulation (TBS) instead of the more widely tested for the CB '1 Hz rTMS' protocol (e.g. section 4.4.9). This was because of the strong encouragement in using TBS protocols as providing more reliable and sizeable results than 1 Hz TMS protocols (Dr. Neil Muggleton; Prof. Vince Walsh, personal communication, 28 January 2009; section 4.4.5). However, the CB TMS literature has only recently exhibited the first TBS findings, and, as already mentioned, they show effects in the opposite direction from those after 1 Hz CB TMS (section 4.4.9). It would thus be interesting to examine whether effects in opposite directions would be established with 1 Hz TMS in the same or similar tasks.

As discussed in chapter 4, TMS experimentation is often based on imaging evidence for the association of a particular brain area with task performance, and goes on to explore the causal nature of such an association (Paus, 1999; see section 4.2 here). Another limiting condition was thus the absence of any MRI coordinates of activated CB loci in tasks similar to those used here. Future research could furthermore combine fMRI to first localize the particular CB area activated in the particular linguistic task for each participant, and then accurately target this area with TMS (e.g. Desmond *et al*., 2005).

Moreover, the NCBV proved to be the most accessible area of stimulation, both because of its significantly smaller depth, as compared with other NCB sites, and because of its longitudinal compactness. The PLCB site employed in the first and second studies (chapters 5 and 6) proved to be a reliable control site. The lack of coils for deeper stimulation and the fact that neuronavigational software was not employed also meant that the NCBV would be the most accessible site. Furthermore, significantly larger amplitudes (45% of maximum machine output instead of the regular 40%) were used in compensation for the depth of the target areas, which involved tolerable yet often unpleasant sensation during stimulation. For the stimulation of deeper CB areas, the double cone coil or the more modern H-coil (e.g. Zangen *et al.*, 2005) would be required. Finally, the use of neuronavigational software in combination with smaller coils for more precise stimulation would help localize the more elusive hemispheric NCB loci ('PLCB', 'Crus I'). Overall, such enhanced TMS agenda would necessitate resources to support longer stays in TMS labs, or, alternatively, the organization of a TMS lab in the University of Edinburgh.

8.2.2. On Patient Studies

A clinical paradigm was avoided here, on the grounds of the significant methodological problems involved in CB cognitive research (section 4.3). Characteristically, while sentence processing tasks had been initially constructed and assessed in normative and pilot studies for both native speakers of English and Modern Greek, the situation in both Edinburgh and Thessaloniki proved discouraging for such patient studies. No sizeable subject pool was available in Hospitals cooperating with the University of Edinburgh (Dr. Thomas Bak, personal communication, 2007-2010). As for the Hospitals in Thessaloniki, despite the availability of patients with CB pathology, several issues made their participation impossible. Patients were constrained to constantly visited hospital beds, and could not be taken to a quiet room for testing. The

great heterogeneity in the etiology of their CB impairment (in most cases accompanied by supratentorial lesions) would have required significant time-depth with each subject, in order for different batteries to be run before the main task. Brain images were also largely unavailable.

In the near future, such patient studies could be arranged in close cooperation with individual neurologists (Dr. K. Gymnopoulos, St. Luke's Hospital, Panorama, Thessaloniki, Greece; personal communication, November, 2009)), ensuring accessible subjects with available brain images in satisfactory testing conditions. Tasks of sentence processing as those initially prepared, with accuracy rates providing the primary dependent measure, would be best suited for such populations (section 4.4.10): e.g. plausibility or θ-role judgments (Ferreira & Stacey, 2000; Ferreira, 2003), or simple encyclopedic question-answering tasks for semantic illusions (Erickson & Mattson, 1981).

8.2.3. On the Psycholinguistic Tasks Involved

Regarding the experimental findings at hand, an outstanding issue would be the direct comparison between formal- and semantic-associative priming, since the studies reported here selectively contrasted types of semantic-categorical (superordinate-subordinate, or coordinate/synonymous-coordinate/synonymous terms) priming with either formal-associative priming (chapters 5 and 6), or with semantic-associative priming (chapter 7). While such a comparison was attempted in the first pilot experiment of the last study (section 7.2), the semantically unrelated, formally co-occurring noun-verb pairs (e.g. 'baby-sitting'), did not show any priming, and their group was thus discarded.

Furthermore, a number of design issues could be raised for the tasks involved. The first experiment (chapter 5) used a small number of items (15 unrelated control items compared with 15 related target items per phase per priming type). Items were not counterbalanced with respect to their relatedness, thus making an items-based analysis of priming sizes impossible. The second study was constrained by the unavailability of any word generation or similarity judgment norms in Modern Greek (section 6.2.1 for discussion). The task of the third TMS study showed priming only in the second phase of the session, suggesting that strategic effects might have been at work. The between-subjects design that was necessarily adopted led to poorly-matched priming sizes in baseline conditions. Moreover, while significantly stronger enhancement of priming was observed in the group that received stimulation of the lateral site, it would have been preferable to show such changes in the absence of a strategic component (section 7.4.7). This could conceivably be achieved with shorter SOAs, or might require another priming paradigm altogether (e.g. naming latencies). Furthermore, the fact that may TMS effects were demonstrated exclusively in either a subjects- or an items-based analysis would suggest that the changes pertained to a subset of the stimuli used, that only a subset of participants were successfully stimulated in the correct NCB sites, and/or that more items and participants were necessary. Future work should employ more homogeneous and well-normed stimulus sets, along with more accurate stimulation methods (section 8.2.1).

Other paradigms for studying simple word-to-word priming phenomena, such as naming latencies tasks, might conceivably induce more robust priming effects, and thus support easier-to-capture TMS-induced changes. However, given the technical issues involved in recording voice responses, the lexical decision task was preferred as a technically less-demanding task (see section 7.1.2). Larger subject pools would once again allow the use of such paradigms.

Furthermore, employing TMS methodology in this new field of NCB emulation-based psycholinguistic mechanisms was a significant factor in using a priming paradigm. Far-reaching conclusions for sentence processing could still be drawn, with the advantage of avoiding any confounds from other constraints and phenomena involved in the sentential level of processing (section 4.4.11). Given that the first encouraging steps were taken here, future research could expand to other experimental hypotheses directly derived from the NCBKFLP model (section 2.6). Above all, it would be interesting to involve sentential stimuli in future CB TMS setups (section 2.6.1). However, the most appropriate sentence comprehension tasks for TMS investigation would be those involving reaction times as a dependent measure, and not accuracy rates (section 4.4.10). Encouragingly, Ward's (2007) studies on phenomena of shallow processing in θ-role- and plausibility- judgment tasks successfully used reaction times as the primary dependent measure, in contrast to those of Ferreira and Stacey (2000) and Ferreira (2003). Similarly, studies involving more subtle judgments, such as noise discrimination in phoneme restoration experiments (section 2.6.3.1), would conceivably be more sensitive to TMS effects, and, as such, they could employ accuracy rates as the dependent measure.

8.2.4. Theoretical Directions

In the same vein, a number of theoretical issues deserve further investigation. For instance, while emulation is involved in covert imitative linguistic perception and comprehension of conspecifics (Pickering & Garrod, 2007), NCB internal models were investigated here only in the context of copying the dynamics of the perception-comprehension process, with no particular reference to the production system. As already discussed, though, there is some first evidence supporting the idea of NCB involvement in covert imitative perception of conspecifics (section 2.5), with which the work here may be fully compatible.

Furthermore, CB computations are involved in both sensory processing and motor output generation (section 2.2). It is by analogy conceivable, then, that the NCBKFLP is not exhausted to language comprehension. Cases of speech reduction based on predictability, as observed in both psycholinguistic (section 2.6.3.2) and historical linguistic contexts (3.6.3.1), provide a suggestive example. Similarly, NCB computations in language processing were only explored on the basis of feedforward internal models; inverse models were not discussed (section 2.2.1.1), but could offer significant insight into further neurocomputational aspects of psycholinguistic mechanisms.

Another interesting research venue would be the investigation of the distinct computations of other neuroanatomical loci, such as the basal ganglia. The fluent gating of competing alternatives for online unification in sentence processing was already seen as a promising linguistic function of that structure. This could be important for the constant negotiation of categorical status of linguistic items, as reflected in grammaticalization changes (section 3.7).

8.3. Closing Statement

The work presented here will hopefully bring together a number of very modern trends of research in cognitive science, such as those of neocerebellar language functions, neocerebellar feedforward control and state estimation in higher cognition, and emulation in language processing. The work has also enriched cerebellar TMS research, by introducing the first linguistic experiments in the field. Finally, the dynamic neurolinguistic components discussed here pave the way for a concrete grounding of the cognitive aspects of major phenomena of historical linguistic change. The thesis then, if anything, has contributed to laying the foundations for studying neocerebellar 'dynamics memory' in language processing, from both a synchronic and a diachronic perspective.

References

Ackermann, H., & Hertrich, I. (2000). The contribution of the cerebellum to speech processing. *Journal of Neurolinguistics, 13,* 95-116.

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., 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.

Aitchison, J. (2003). Psycholinguistic perspectives on language change. In B. Joseph & R. Janda (Eds.), *Handbook of historical linguistics* (pp. 736-743). Oxford: Blackwell.

Al-Anazi, A., Hassounah, M., Sheikh, B., & Barayan, S. (2001). Cerebellar mutism caused by arteriovenous malformation of the vermis. *British Journal of Neurosurgery, 15*(1), 47-50.

Amassian, V. E., Cracco, R. Q., Maccabee, P. J., & Cracco, J. B. (1992). Cerebellofrontal cortical projections in humans studied with the magnetic coil. *Electroencephalography and Clinical Neurophysiology, 85,* 265-272.

Anand, S., & Hotson, J. (2002). Transcranial magnetic stimulation: neurophysiological applications and safety. *Brain and Cognition, 50,* 366-386.

Andreasen, N. C. (1986). Scale for the assessment of Thought, Language, and Communication (TLC). *Schizophrenia Bulletin, 12,* 473-482.

Antilla, R. (1989 [1972]). *Historical and comparative linguistics* (2nd ed.) Amsterdam: John Benjamins.

Arbib, M. A., Érdi, P., & Szentágothai, J. (1998). *Neural organization: Structure, function, and dynamics.* Cambridge, Massachusetts: MIT Press.

Argyropoulos, G. P. (2008a). 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. (2008b, June). *Grammaticalization and the corticocerebellar teleonomy.* Paper presented at the 'Cognitive and Functional Perspectives on Dynamic Tendencies in Languages' conference, University of Tartu, Estonia.

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. (2010a). Is grammaticalization glossogenetic? In A. D. M. Smith, M. Schouwstra, B. de Boer, & K. Smith (Eds.), *The evolution of language: Proceedings of the 8th international conference on the evolution of language* (pp. 3-10). Singapore: World Scientific.

Argyropoulos, G. P. (2010b). *Theta-burst stimulation of the right neocerebellar vermis selectively enhances lexical associative priming.* Manuscript submitted for publication.

Argyropoulos, G. P., Kimiskidis, V., & Papagiannopoulos, S. (2010). *Theta-burst stimulation of the right neocerebellar vermis selectively disrupts reading processes.* Manuscript in preparation.

Argyropoulos, G. P., & Muggleton, N. (2010). *Effects of neocerebellar theta-burst stimulation on thematic role- based priming.* Manuscript in preparation.

Aylett, M., & Turk, A. (2004). The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Language and Speech*, *47*(1), 31-56.

Bard, E. G., Anderson, A. H., Sotillo, C., Aylett, M., Doherty-Sneddon, G., & Newlands, A. (2000). Controlling the intelligibility of referring expressions in dialogue. *Journal of Memory and Language*, *42* (1), 1-22.

Barlow, J. S. (2002). *The cerebellum and adaptive control*. Cambridge: Cambridge University Press.

Barton, S., & Sanford, A. J. (1993). A case study of anomaly detection: Shallow semantic processing and cohesion establishment. *Memory & Cognition*, *21*, 477-487.

Bastian, A. J., Martin, T. A., Keating, J. G., & Thach, W. T. (1996). Cerebellar ataxia: Abnormal control of interaction torques across multiple joints. *Journal of Neurophysiology*, *76*, 492-509.

Batali, J. (2002). The negotiation and acquisition of recursive grammars as a result of competition among exemplars. In E. Briscoe (Ed.), *Linguistic evolution through language acquisition: Formal and computational models* (pp. 111-172). Cambridge: Cambridge University Press.

Bauer, L. (1983). *English word-formation.* Cambridge: Cambridge University Press.

Becker, T. (1993). Back-formation, cross-formation, and 'bracketing paradoxes' in paradigmatic morphology. In G. Booij & J. van Marle (Eds.), *Yearbook of morphology 1993* (pp. 1-25). Dordrecht: Kluwer.

Bellugi, U., Bihrle B., Jernigan, T., Trauner, D., & Doherty, S. (1990). Neuropsychological, neurological, and neuroanatomical profile of Williams syndrome. *American Journal of Medical Genetics, 37*(S1), 115-125.

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,* 1879-1884.

Blakemore, S.- J., Wolpert, D. M., & Frith, C. D. (1998). Central cancellation of self-produced tickle sensation. *Nature Neuroscience, 1*(7), 635-640.

Blakeslee, S. (1994, November 8). Theory on human brain hints how its unique traits arose. *The New York Times: Science.* Retrieved from http://www.nytimes.com/1994/11/08/science/theory-on-human-brain-hints-how-its-unique-traits-arose.html

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

Botez, M. I., Botez, T., Elie, R., & Attig, E. (1989). Role of the cerebellum in complex human behavior. *Italian Journal of Neurological Science, 10,* 291-300.

Botez, M. I., Gravel, J., Attig, E., & Vézina, J.- L. (1985). Reversible chronic ataxia after phenytoin intoxication: Possible role of the cerebellum in cognitive thought. *Neurology, 35,* 1152-1157.

Bouzouita, M. (2002). *Clitic placement in Old and Modern Spanish: A dynamic account.* Unpublished Master's thesis. King's College London.

Boye, K. (2008, August). *Grammatical expressions and grammaticalization - a functional theory.* Paper presented at the annual Salos Linguistics Summer School

in Litauen, Lithuania [PowerPoint slides]. Retrieved from http://academiasalensis.org/club_rtm_v1_vb/salos2008/Boye1.pdf

Boye, K., & Harder, P. (2009). Evidentiality. Linguistic categories and grammaticalization. *Functions of Language*, *16*(1), 9-43.

Boyland, J. T. (1996). *Morphosyntactic change in progress: A psycholinguistic approach.* Unpublished doctoral dissertation, University of California, Berkeley.

Bracke-Tolkmitt, R., Linden, A., Canavan, A. G. M., Rockstroch, B., Scholz, E., Wessel, K., & Diener, H.- C. (1989). The cerebellum contributes to mental skills. *Behavioral Neuroscience*, *103*(20), 442-446.

Bredart, S., & Modolo, K. (1988). Moses strikes again: Focalization effects on a semantic illusion. *Acta Psychologica*, 67, 135-144.

Brighina, F., Daniele, O., Piazza, A., Giglia, G., & Fierro, B. (2006). Hemispheric cerebellar rTMS to treat drug-resistant epilepsy: case reports. *Neuroscience Letters*, *397*, 229-33.

Brighton, H., Kirby, S., & Smith, K. (2005). Cultural selection for learnability: Three hypotheses underlying the view that language adapts to be learnable. In M. Tallerman (Ed.), *Language origins: Perspectives on evolution* (pp. 291-309). Oxford: Oxford University Press.

Brindley, G. S. (1964). The use made by the cerebellum of the information that it receives from sense organs (report on symposia and meetings). *International Brain Research Organization Bulletin*, *3*(30), 80.

Briscoe, E. (Ed.) (2002). *Linguistic evolution through language acquisition: Formal and computational models.* Cambridge: Cambridge University Press.

Brodal, A. (1981). *Neurological anatomy in relation to clinical medicine.* Oxford: Oxford University Press.

Broich, K., Hartmann, A., Biersck, H. J., & Horn, R. (1987). Crossed cerebellocerebral diaschisis in a patient with cerebellar infarction. *Neuroscience Letters*, *83*, 7-12.

Brown, S. M., Kieffaber, P. D., Carroll, C. A., Vohs, J. L., Tracy, J. A., Shekhar, A., O' Donnell, B. F., Steinmetz, J. E., & Hetrick, W. P. (2005). Eyeblink conditioning deficits indicate timing and cerebellar abnormalities in schizophrenia. *Brain and Cognition*, *58*, 94-108.

Bybee, J. (1995). Regular morphology and the lexicon. *Language and Cognitive Processes*, *10*, 425-455.

Bybee, J. (1998). A functionalist approach to grammar and its evolution. *Evolution of Communication*, *2*, 249-278.

Bybee, J. (2002a). Cognitive processes in grammaticalization. In M. Tomasello (Ed.), *The new psychology of language, Vol. II*. (145-168). New Jersey: Erlbaum.

Bybee, J. (2002b). Phonological evidence for exemplar storage of multiword sequences. *Studies in Second Language Acquisition*, *24*, 215-221.

Bybee, J. (2002c). Sequentiality as the basis of constituent structure. In T. Givón & B. Malle (Eds.), *The evolution of language out of pre-language* (pp. 107-132). Amsterdam: John Benjamins.

Bybee, J. (2002d). Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change*, *14*, 261-290.

Bybee, J. (2003). Mechanisms of change in grammaticization: The role of frequency. In B. D. Joseph & J. Janda (Eds.), *The handbook of historical linguistics* (pp. 602-623). Oxford: Blackwell.

Bybee, J. (2006). From usage to grammar: the mind's response to repetition. *Language*, *82*(4), 711-733.

Bybee, J. (2007a). *Frequency of use and the organization of language*. Oxford: Oxford University Press.

Bybee, J. (2007b). Diachronic Linguistics. In D. Geeraerts & H. Cuyckens (Eds.), *The Oxford handbook of cognitive linguistics* (pp. 945-987). Oxford: Oxford University Press.

Bybee, J., Perkins, R., & Pagliuca, W. (1994). *The evolution of grammar. Tense, aspect, and modality in the languages of the world.* Chicago: University of Chicago Press.

Bybee, J., & Thompson, S. (2000). Three frequency effects in syntax. *Berkeley Linguistics Society*, *23*, 65-85.

Campbell, L. (2001). What's wrong with grammaticalization? *Language Sciences*, *23*, 113-161.

Cañas, J. J., & Bajo, M. T. (1994). Strategic associative priming in the lexical decision task. *The Quarterly Journal of Experimental Psychology*, *47*A, 383-405.

Canavan, A. G., Sprengelmeyer, R., Diener, H. C., & Hömberg, V. (1994). Conditional associative learning is impaired in cerebellar disease in humans. *Behavioral Neuroscience, 108*, 475-485.

Caramazza, A., & Zurif, E. B. (1976). Dissociation of algorithmic and heuristic processes in language comprehension: Evidence from aphasia. *Brain and Language, 3*, 572-582.

Carpenter, M. B. (1991). The cerebellum. In T. M. Tracy (Ed.), *Core Text of Neuroanatomy* (pp. 225-247). London: Williams and Wilkins.

Chen, R., Classen, J., Gerloff, C., Celnik, P., Wassermann, E. M., Hallett, M., & Cohen, L. G. (1997). Depression of motor cortex excitability by low-frequency transcranial magnetic stimulation, *Neurology, 48*, 1398-1403.

Chomsky, N. (1965). *Aspects of the theory of syntax.* MIT Press.

Chomsky, N. (1981a). Principles and parameters in syntactic theory. In N. Hornstein & D. Lightfoot (Eds.), *Explanations in linguistics: The logical problem of language acquisition* (pp. 32-75). London: Longman.

Chomsky, N. (1981b). *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*, 526-537.

Christiansen, M. H. (1995). *Language as an organism - implications for the evolution and acquisition of language.* Unpublished manuscript, Washington University.

Christiansen, M. H., & Chater, N. (2008). Language as shaped by the brain. *Behavioral and Brain Sciences, 31*(5), 489-509.

Christiansen, M. H., Kelly, L., Shillcock, R., & Greenfield, K. (in press). Impaired artificial grammar learning in agrammatism. *Cognition.*

Clark, R., & Roberts, I. (1993). A computational model of language learnability and language change. *Linguistic Inquiry, 24*, 299-345.

Cleese, J. (Speaker). (1973). Word association [Audio Recording]. In: Monty Python (CD-ROM), *Free record given away with the Monty Python matching tie and handkerchief* (Side 1, Track 4). Charisma Records/ Arista Records.

Connor, L. T., DeShazo, Braby, T., Snyder, A. Z., Lewis, C., Blasi, V., & Corbetta, M. (2006). Cerebellar activity switches hemispheres with cerebral recovery in aphasia. *Neuropsychologia, 44*, 171-177.

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

Craik, K. (1943). *The Nature of explanation*. Cambridge: Cambridge University Press.

Croft, W. (1995). Intonation units and grammatical structure. *Linguistics*, *33*, 839-882.

Croft, W. (2000). *Explaining language change: An evolutionary approach*. Harlow: Longman.

Croft, W. (2001). *Radical construction grammar*. Oxford: Oxford University Press.

Croft, W. (2004). Form, meaning and speakers in the evolution of language. *Studies in Language*, *28*(3), 608-611.

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. (1997). Neuropsychological abnormalities in cerebellar syndromes- fact or fiction? *International Review of Neurobiology*, *41*, 455-471.

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), 411-419.

Daum, I., Snitz, B. E., & Ackermann, H. (2001). Neuropsychological deficits in cerebellar syndromes. *International Review of Psychiatry*, *13*, 268-275.

de Vries, M. H., Barth, A. C., Maiworm, S., Knecht, S., Zwitserlood, P., & Flöel, A. (2010). Electrical Stimulation of Broca's Area Enhances Implicit Learning of an Artificial Grammar. *Journal of Cognitive Neuroscience*, *22*(11), 2427-2436.

Deacon, T. (1997). *The symbolic species: The coevolution of language and the brain.* New York: Norton.

Decety, J., Grezes, J., Costes, N., Perani, D., Jeannerod, M., Procyk, E., Grassi, F., & Fazio, F. (1997). Brain activity during observation of actions. Influence of action content and subject's strategy. *Brain*, *120*, 1763-1777.

Del Olmo, M.F., Cheeran, B., Koch, G., & Rothwell, J. C. (2007). Role of the cerebellum in externally paced rhythmic finger movements. *Journal of Neurophysiology*, *98*(1), 145-52.

Desmond, J. E., Chen, S. H., & Shieh, P. B. (2005). Cerebellar transcranial magnetic stimulation impairs verbal working memory. *Annals of Neurology, 58*, 553-560.

Desmond, J. E., & Fiez, J. A. (1998). Neuroimaging studies of the cerebellum: Language, learning, and memory. *Trends in Cognitive Sciences*, *2*(9), 355-362.

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*, 368-376.

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.

Diessel, H. (2004). *The acquisition of complex sentences*. Cambridge: Cambridge University Press.

di Lazzaro, V., Pilato, F., Saturno, E., Oliviero, A., Dileone, M., Mazzone, P., Insola, A., Tonali, P. A., Ranieri, F., Huang, Y. Z., & Rothwell, J. C. (2005). Theta-burst repetitive transcranial magnetic stimulation suppresses specific excitatory circuits in the human motor cortex. *Journal of Physiology, 565*, 945-950.

Dow, R. S. (1942). The evolution and anatomy of the cerebellum. *Biological Reviews of the Cambridge Philosophical Society, 17*, 179-220.

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

Doya, K. (1999). What are the computations of the cerebellum, the basal ganglia and the cerebral cortex? *Neural Networks, 12*, 961-974.

Doyon, J., Laforce, R., Bouchard, G., Gaudreau, D., Roy, J., Poirer, M., Bedard, P. J., Bedard, F., & Bouchard, J. P. (1998). Role of the striatum, cerebellum, and frontal lobes in the automatization of a repeated visuomotor sequence of movements. *Neuropsychologica, 36*, 625-641.

Doyon, J., Owen, A. M., Petrides, M., Sziklas, V., & Evans, A. C. (1996). Functional anatomy of visuomotor skill learning in human subjects examined with positron emission tomography. *European Journal of Neuroscience, 8*, 637-648.

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.

Dum, R. P., & Strick, P. L. (2003). An unfolded map of the cerebellar dentate nucleus and its projections to the cerebral cortex. *Journal of Neurophysiology*, *89*, 634-639.

Eccles, J. C., Ito, M., & Szentagothai, J. (1967). *The cerebellum as a neuronal machine*. New York and Heidelberg: Springer-Verlag.

Erickson, T. A., & Mattson, M. E. (1981). From words to meaning: a semantic illusion. *Journal of Verbal Learning: Verbal Behaviour*, *20*, 540-552.

Fabbro, F. (2000). Introduction to language and cerebellum. *Journal of Neurolinguistics*, *13*, 83-94.

Fabbro, F., Moretti, R., & Bava, A. (2000). Language impairments in patients with cerebellar lesions. *Journal of Neurolinguistics*, *13*, 173-188.

Fabbro, F., Tavano, A., Corti, S., Bresolin, N., De Fabritiis, P., & Borgatti, R. (2004). Long-term neuropsychological deficits after cerebellar infarctions in two young adult twins. *Neuropsychologia*, *42*, 536-545.

Faraday, M. (1965). Effects on the production of electricity from magnetism (1831). In L. P. Williams (Ed.), *Michael Faraday* (p. 531). New York: Basic Books.

Fawcett A. J., Nicolson, R. I. (1999). Performance of dyslexic children on cognitive and cerebellar tests. *Journal of Motor Behavior*, *31*, 68-78.

Fawcett, A. J., Nicolson, R. I., Dean, P. (1996). Impaired performance of children with dyslexia on a range of cerebellar tasks. *Annals of Dyslexia*, *46*, 259-83.

Fellbaum, C. (Ed.) (1998). *WordNet: An electronic lexical database*. MIT Press.

Ferrand, L., & New, B. (2003). Semantic and associative priming in the mental lexicon. In P. Bonin (Ed.), *Mental lexicon: Some words to talk about words* (pp. 25-43). Hauppauge, New York: Nova Science.

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

Ferreira, F., Bailey, K. G. D., & Ferraro, V. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science*, *11*, 11-15.

Ferreira, F., & Stacey, J. L. (2000). *The misinterpretation of passive sentences.* Unpublished manuscript, Michigan State University.

Ferretti, T. R., Gagné, C., & McRae, K. (2003). Thematic role focusing by participle inflections: Evidence from conceptual combination. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*(1), 118-127.

Ferretti, T. R., Kutas, M., & McRae, K. (2007). Verb aspect and the activation of event knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(1), 182-196.

Ferretti, T. R., McRae, K., & Hatherell, A. (2001). Integrating verbs, situation schemas, and thematic role concepts. *Journal of Memory and Language*, *44*, 516-547.

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), 1687-1697.

Fierro, B., Giglia, G., Palermo, A., Pecoraro, C., Scalia, S., & Brighina, F. (2007a). Modulatory effects of 1 Hz rTMS over the cerebellum on motor cortex excitability. *Experimental Brain Research*, *176*(3), 440-447.

Fierro, B., Palermo, A., Puma, A., Francolini, M., Panetta, M., Daniele, O., & Brighina, F. (2007b). Role of the cerebellum in time perception: A TMS study in normal subjects. *Journal of the Neurological Sciences*, *263*(1), 107-112.

Fiez, J. A., Petersen, S. E., Cheney, M. K., & Raichle, M. E. (1992). Impaired non-motor learning and error detection associated with cerebellar damage. A single case study, *Brain*, *115*, 155-178.

Fiez, J. A., & Raichle, M. (1997). Linguistic processing. In J. D. Schmahmann (Ed.), R. J. Bradley, R. A. Harris, & P. Jenner (Series Eds.), *The cerebellum and cognition. International Review of Neurobiology, 41* (pp. 233-54). San Diego: Academic Press.

Fillenbaum, S. (1974). Pragmatic normalization: Further results for some conjunctive and disjunctive sentences. *Journal of Experimental Psychology*, *102*, 574-578.

Finch, A. J., Nicolson, R. I., & Fawcett, A. J. (2002). Evidence for a neuroanatomical difference within the olivo-cerebellar pathway of adults with dyslexia. *Cortex*, *38*, 529-539.

Fleischman, S. (1982). *The future in thought and language: Diachronic evidence from Romance*. Cambridge: Cambridge University Press.

Forkstam, C., Hagoort, P., Fernández, G., Ingvar, M., & Petersson, K. M. (2006). Neural correlates of artificial syntactic structure classification. *NeuroImage, 32*, 956-967.

Forrest, D. (1976). Nonsense and sense in schizophrenic language. *Schizophrenia Bulletin, 2*, 286-298.

Forster, K. I., & Davis, C. (1991). The density constraint on form-priming in the naming task: Interference from a masked prime. *Journal of Memory and Language, 30*, 1-25.

Forster, K. I., & Olbrei, I. (1973). Semantic heuristics and syntactic analysis. *Cognition, 2*, 319-347.

Frank, B., Schoch, B., Richter, S., Frings, M., Karnath, H. O., & Timmann, D. (2007). Cerebellar lesion studies of cognitive function in children and adolescents - limitations and negative findings. *The Cerebellum, 6*, 242-253.

Frazier, L. & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology, 14*, 178-210.

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.

Garnham, A., & Oakhill, J. V. (1996). The mental models theory of language comprehension. In B. K. Britton & A. C. Graesser (Eds.), *Models of understanding text* (pp. 313-339). Hillsdale, New Jersey: Erlbaum.

Garrett, M. F. (2000). Remarks on the architecture of language processing systems. In Y. Grodzinsky & L. P. Shapiro (Eds.), *Language and the brain: Representation and processing. Foundations of neuropsychology series* (pp. 31-69). San Diego, CA: Academic Press.

Garrod, S., & Anderson, A. (1987). Saying what you mean in dialogue: a study in conceptual and semantic co-ordination. *Cognition, 27*,181-218.

Garrod, S., & Doherty, G. (1994). Conversation, co-ordination and convention: An empirical investigation of how groups establish linguistic conventions. *Cognition, 53*, 181-215.

Garrod, S., & Pickering, M. J. (2004). Why is conversation so easy? *Trends in Cognitive Sciences, 8*, 8-11.

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*, 318-333.

Gelb, A. (1974). *Applied optimal control*. Cambridge, Massachusetts: MIT Press.

Gentner, R., Wankerl, K., Reinsberger, C., Zeller, D., & Classen, J. (2008). Depression of human corticospinal excitability induced by magnetic theta-burst stimulation: Evidence of rapid polarity-reversing metaplasticity. *Cerebral Cortex, 18*(9), 2046-2053.

Gernsbacher, M. A., Tallent, K. A., & Bolliger, C. M. (1999). Disordered discourse in schizophrenia described by the structure building framework. *Discourse Studies, 1*, 355-372.

Givón, T. (1975). Serial verbs and syntactic change: Niger-Congo. In C. N. Li (Ed.), *Word order and word order change* (pp. 47-112). Austin: University of Texas Press.

Givón, T. (1979). *On understanding grammar*. New York: Academic Press.

Givón, T. (1989). *Mind, code and context: Essays in pragmatics*. New York: Academic Press.

Givón, T. (1991). Serial verbs and the mental reality of 'event': grammatical versus cognitive packaging. In E. C. Traugott & B. Heine (Eds.), *Approaches to grammaticalization, Vol. 1*, (pp. 81-127). Amsterdam: John Benjamins.

Givón, T. (1995). *Functionalism and grammar*. Amsterdam: John Benjamins.

Givón, T. (2009). *The genesis of syntactic complexity*. Amsterdam: John Benjamins.

Glenberg, A. M. (2009). Prediction and emotion in dialogue. *European Journal of Social Psychology, 39*, 1169-1172.

Goldberg, A. E. (2003). Constructions: a new theoretical approach to language. *Trends in Cognitive Sciences, 7*(5), 219-224.

Goldman-Rakic, P. S. (1995). Cellular basis of working memory. *Neuron, 14*, 477-485.

Goodkin, H. P., & Thach, W. T. (2003). Cerebellar control of constrained and unconstrained movements. I. Nuclear inactivation. *Journal of Neurophysiology, 89*, 884-895.

Grafman, J., Litvan, I., Massaquoi, S., Stewart, M., Sirigu, A., & Hallett, M. (1992). Cognitive planning deficit in patients with cerebellar atrophy. *Neurology, 42*, 1493-1496.

Grodzinsky, Y., & Finkel, L. (1998). The neurology of empty categories. Aphasics' failure to detect ungrammaticality. *Journal of Cognitive Neuroscience, 10*(2), 281-292.

Grossman, M., Lee, C., Morris, J., Stern, M. B., & Hurtig, H. I. (2002). Assessing resource demands during sentence processing in Parkinson's disease. *Brain and Language, 80*, 603-616.

Grush, R. (1996). *Emulation and cognition.* Unpublished doctoral dissertation, University of California, San Diego.

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

Hagoort, P. (1997). Semantic priming in Broca's aphasics at a short SOA: No support for an automatic access deficit. *Brain and Language, 56*, 287-300.

Hagoort, P. (2003). How the brain solves the binding problem for language: a neurocomputational model of syntactic processing. *NeuroImage, 20*, S18-S29.

Hagoort, P. (2005). On Broca, brain, and binding: a new framework. *Trends in Cognitive Sciences, 9*, 416-423.

Haiman, J. (1994). Ritualization and the development of language. In W. Pagliuca (Ed.), *Perspectives on Grammaticalization* (pp. 3-28). Amsterdam: John Benjamins.

Hare, M., & Elman, J. L. (1995). Learning and morphological change. *Cognition, 56*, 61-98.

Harper, J. W., & Heath, R. G. (1973). Anatomic connections of the fastigial nucleus to the rostral forebrain in the cat. *Experimental Neurology, 39*, 285-292.

Harris, A., & Campbell, L. (1995). *Historical syntax in cross-linguistic perspective*. Cambridge: Cambridge University Press.

Hartelius, L., Runmarker, B., Andersen, O., & Nord, L. (2000). Temporal speech characteristics of individuals with multiple sclerosis and ataxic dysarthria: 'Scanning speech' revisited. *Folia Phoniatrica et Logopaedica, 52*, 228-238.

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

Hashimoto, T., & Nakatsuka, M. (2006). Reconsidering Kirby's compositionality model toward modeling grammaticalization. In A. Cangelosi, A. D. M. Smith, & K. Smith (Eds.), *The evolution of language, proceedings of the 6th international conference* (pp. 415-416). Singapore: World Scientific.

Haspelmath, M. (1998). Does grammaticalization need reanalysis? *Studies in Language, 22*, 315-51.

Haspelmath, M. (1999). Why is grammaticalization irreversible? *Linguistics, 37(6)*, 1043-1068.

Haspelmath, M. (2004). On directionality in language change with particular reference to grammaticalization. In O. Fischer, M. Norde, & H. Perridon (Eds.), *Up and down the cline: The nature of grammaticalization* (pp. 17-44). Amsterdam: John Benjamins.

Haspelmath, M. (2008). Reply to Haiman and Croft. *Cognitive Linguistics, 19*(1), 59-66.

Hatzigeorgiu, N., Gavrilidou, M., Piperidis, S., Carayannis, G., Papakostopoulou, A., Spiliotopoulou, A., Vacalopoulou, A., Labropoulou, P., Mantzari, E., Papageorgiou, H., & Demiros I. (2000). Design and implementation of the online ILSP Greek Corpus. *Proceedings of the Second International Conference of Language Resources and Evaluation (LREC), 3*, 1737-1740.

Hazeltine, E. & Ivry, R. B. (2002). Can we teach the cerebellum new tricks? *Science, 296*, 1979-1980.

Hawkins, J. A. (2004). *Efficiency and complexity in grammars*. Oxford: Oxford University Press.

Heine, B. (2003). Grammaticalization. In B. D. Joseph & R. D. Janda (Eds.), *The handbook of historical linguistics* (pp. 575-601). Oxford: Blackwell.

Heine, B. & Kuteva, T. (2002). *World lexicon of grammaticalization.* Cambridge: Cambridge University Press.

Heine, B., & Reh, M. (1984). *Grammaticalization and reanalysis in African languages*. Hamburg: Buske.

Helmuth, L. L., Ivry, R. B., & Shimizu, N. (1997). Preserved performance by cerebellar patients on tests of word generation, discrimination learning, and attention. *Learning and Memory, 3*, 456-474.

Himmelmann, N. P. (2004). Lexicalization and grammaticalization: Opposite or orthogonal? In W. Bisang, N. P. Himmelmann, & B. Wiemer (Eds.), *What makes grammaticalization? A look from its fringes and its components* (pp. 21-42). Berlin: Mouton de Gruyter.

Hochstadt, J., Nakano, H., Lieberman, P., & Friedman, J. (2006). The roles of sequencing and verbal working memory in sentence comprehension deficits in Parkinson's disease. *Brain and Language, 97*, 243-225.

Hoefler, S. & Smith, A. D. M. (2008). Reanalysis vs. metaphor? What grammaticalisation can tell us about language evolution. 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. 163-170). Singapore: World Scientific.

Hoefler, S., & Smith, A. D. M. (2009). The pre-linguistic basis of grammaticalisation: A unified approach to metaphor and reanalysis. *Studies in Language, 33*(4), 883-906.

Holmes, G. (1917). The symptoms of acute cerebellar injuries due to gunshot injuries. *Brain, 40*, 461-535.

Holmes, G. (1939). The cerebellum of man. *Brain, 62*, 1-30.

Hooper, J. B. (1976). Word frequency in lexical diffusion and the source of morphophonological change. In W. Christie (Ed.), *Current progress in historical linguistics* (pp. 95-105). Amsterdam: North Holland.

Hopper, P., & Traugott, E. C. (1993). *Grammaticalization*. Cambridge: Cambridge University Press.

Hossain, M. A., Russell, J. C., Miknyoczki, S., Ruggeri, B., Lal, B., & Laterra, J. (2004). Vascular endothelial growth factor mediates vasogenic edema in acute lead encephalopathy. *Annals of Neurology, 55*, 660-667.

Houk, J. C., & Wise, S. P. (1995). Distributed modular architectures linking basal ganglia, cerebellum, and cerebral cortex: Their role in planning and controlling action. *Cerebral Cortex, 2*, 95-110.

Huang, Y.- Z., Edwards, M. J., Rounis, E., Bhatia, K. P., & Rothwell, J. C. (2005). Theta burst stimulation of the human motor cortex. *Neuron, 45*, 201-206.

Hubbard, E. M., & Ramachandran, V. S. (2004). The size-weight illusion, emulation, and the cerebellum. *Behavioral and Brain Sciences, 27*, 407-408. Commentary on R. Grush, *The emulation theory of representation: motor control, imagery, and perception.*

Hurford, J. R. (1987). *Language and number: The emergence of a cognitive system*. Oxford: Blackwell.

Hurford, J. R. (1991). Nativist and functional explanations in language acquisition. In I. Roca (Ed.), *Logical issues in language acquisition* (pp. 85-136). Dordrecht: Holland Foris.

Hurford, J. R. (2002). Expression/induction models of language evolution: dimensions and issues. In T. Briscoe (Ed.), *Linguistic evolution through language acquisition: formal and computational models* (pp. 301-344). Cambridge: Cambridge University Press.

Hurford, J. R. (2009). Universals and the diachronic life cycle of languages. In M. Christiansen, C. Collins, & S. Edelman (Eds.), *Language universals* (pp. 40-53). Oxford, New York: Oxford University Press.

Hurford, J. R., Flaherty, M., & Argyropoulos, G. (2007). Past and future, human and non-human, semantic/procedural and episodic. *Behavioral and Brain Sciences, 30*(3), 324-325. Commentary on T. Suddendorf & M. Corballis *The evolution of foresight: What is mental time travel and is it unique to humans?*

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. (2005). Understanding others: Imitation, language, empathy. In: S. Hurley & N. Chater (Eds.), *Perspectives on imitation: From cognitive neuroscience to social science, Volume 1: mechanisms of imitation and imitation in animals* (pp. 77- 99). Cambridge, Massachusetts: MIT Press.

Iacoboni, M. (2008). The role of premotor cortex in speech perception: Evidence from fMRI and rTMS. *Journal of Physiology, 102*(1-3), 31-34.

Ichimiya, T., Okubo, Y., Suhara, T., & Sudo, Y. (2001). Reduced volume of the cerebellar vermis in neuroleptic-naïve schizophrenia. *Biological Psychiatry, 49,* 20-27.

Imamizu, H., Kuroda, T., Miyauchi, S., Yoshioka, T., & Kawato, M. (2003). Modular organization of internal models of tools in the human cerebellum. *Proceedings of*

the National Academy of Sciences of the United States of America, *100*(9), 5361-5466.

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*, 192-195.

Ito, M. (1970). Neurophysiological basis of the cerebellar motor control system. *International Journal of Neurology*, *7*, 162-176.

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

Ito, M. (1990a). A new physiological concept on cerebellum. *Revue Neurologique (Paris)*, *146*, 564-569.

Ito, M. (1990b). Neural control as a major aspect of high-order brain function. In J. C. Eccles & O. Creutzfeldt (Eds.), *The principles of design and operation of the brain (Experimental Brain Research Supplement)* (pp. 281-292). Berlin: Springer-Verlag

Ito, M. (1993a). Movement and thought: Identical control mechanisms by the cerebellum. *Trends in Neurosciences*, *16*(11), 448-450.

Ito, M. (1993b). New concepts in cerebellar function. *Revue Neurologique*, *149*(11), 596-599.

Ito, M. (1997). Cerebellar microcomplexes. In J. D. Schmahmann (Ed.), R. J. Bradley, R. A. Harris, & P. Jenner (Series Eds.), *The cerebellum and cognition. International Review of Neurobiology*, *41* (pp. 475-487). San Diego: Academic Press.

Ito, M. (2000). Neural control of cognition and language. In A. Marantz, Y. Miyashita, & W. O'Neil (Eds.), *Image, language, brain* (pp. 149-162). Cambridge, Massachusetts: MIT Press.

Ito, M. (2002). Historical review of the significance of the cerebellum and the role of Purkinje cells in motor learning. *Annals of the New York Academy of Sciences*, *978*, 273-288.

Ito, M. (2008). Control of mental activities by internal models in the cerebellum. *Nature Reviews Neuroscience*, *9*, 304-313.

Ito, M., Sakurai, M., & Tongroach, P. (1982). Climbing fiber induced depression of both mossy fiber responsiveness and glutamate sensitivity of cerebellar Purkinje cells. *Journal of Physiology (London)*, *324*, 113-134.

Ivry, R. B., & Fiez, J. A. (2000). Cerebellar contributions to cognition and imagery. In M. S. Gazzaniga (Ed.), *The new cognitive neurosciences* (pp. 999-1011). Cambridge, MA: MIT Press.

Ivry, R. B., & Justus, T. C. (2001). A neural instantiation of the motor theory of speech perception. *Trends in Neurosciences, 24*(9), 513-515.

Jackendoff, R. (2002). *Foundations of language: Brain, meaning, grammar, evolution.* Oxford: Oxford University Press.

Jahanshahi, M., & Rothwell, J. (2000). Transcranial magnetic stimulation studies of cognition: an emerging field. *Experimental Brain Research, 131*, 1-9.

Jarvis, B. G. (2008). *DirectRT* (version 2008.1.0.11) [Computer software]. New York: Empirisoft Corporation.

Jenkins, I. H., Brooks, D. J., Nixon, P. D., Frackowiak, R. S. J., & Passingham, R. E. (1994). Motor sequence learning: a study with positron emission tomography. *Journal of Neuroscience, 14*, 3775-3790.

Jordan, M. I., & Rumelhart, D. E. (1992). Forward models: Supervised learning with a distal teacher. *Cognitive Science, 16*, 307-354.

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*, 1325-1337.

Jueptner, M., & Weiller, C. (1998). A review of differences between basal ganglia and cerebellar control of movements as revealed by functional imaging studies. *Brain, 121*, 1437-1449.

Jurafsky, D., Bell, A., Gregory, M., & Raymond, W. D. (2001). Probabilistic relations between words: Evidence from reduction in lexical production. In J. Bybee & P. Hopper (Eds.), *Frequency and the emergence of linguistic structure* (pp. 229-254). Amsterdam: John Benjamins.

Justus, T. (2004). The cerebellum and English grammatical morphology: evidence from production, comprehension, and grammaticality judgments. *Journal of Cognitive Neuroscience, 16*(7), 1115-1130.

Justus, T., & Ivry, R. B. (2001). The cognitive neuropsychology of the cerebellum. *International Review of Psychiatry*, *13*, 276-282.

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, *82*, 35-45.

Kalman, R. E., & Bucy, R. S. (1961). New results in linear filtering and prediction. *Journal of Basic Engineering*, *83D*, 95-108.

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

Kawato, M., Furukawa, K., & Suzuki, R. (1987). A hierarchical neural-network model for control and learning of voluntary movement. *Biological Cybernetics*, *57*, 169-185.

Kawato, M., & Gomi, H. (1992). A computational model of four regions of the cerebellum based on feedback-error learning. *Biological Cybernetics*, *68*, 95-103.

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*, 171-188.

Keesing, R. M. (1991). Substrates, calquing and grammaticalization in Melanesian Pidgin. In E. C. Traugott & B. Heine (Eds.), *Approaches to grammaticalization*. *Vol. 1* (pp. 315-42). Amsterdam and Philadelphia: John Benjamins.

Kempson, R., & Cann, R. (2007). Dynamic Syntax and dialogue modelling: preliminaries for a dialogue-driven account of syntactic change. In J. Salmons & S. A. Dubenion-Smith (guest Eds.), *Historical Linguistics 2005* (*Current issues in linguistic theory*, *284*) (pp. 73-101) Amsterdam & Philadelphia: John Benjamins

Kirby, S. (1999). *Function, selection and innateness: The emergence of language universals*. Oxford: Oxford University Press.

Kirby, S. (2001). Spontaneous evolution of linguistic structure: An iterated learning model of the emergence of regularity and irregularity. *IEEE Journal of Evolutionary Computation*, *5*(2), 102-110.

Kirby, S. (2002). Learning, bottlenecks, and the evolution of recursive syntax. In E. J. Briscoe (Ed.), *Linguistic evolution through language acquisition* (pp. 173-203). Cambridge: Cambridge University Press.

Kirby, S., & Hurford, J. (1997). Learning, culture and evolution in the origin of linguistic constraints. In P. Husbands & I. Harvey (Eds.), *European Conference on Artificial Life 97*, (pp. 493-502). MIT: MIT Press.

Kirby, S., Smith, K., & Brighton, H. (2004). From UG to universals: Linguistic adaptation through iterated learning. *Studies in Language*, *28*(3), 587-607.

Kiss, G. R., Armstrong, C., Milroy, R., & Piper, J. (1973). An associative thesaurus of English and its computer analysis. In A. J. Aitken, R. W. Bailey, & N. Hamilton-Smith (Eds.), *The computer and literary studies* (pp. 153-165). Edinburgh: Edinburgh University Press.

Knowlton B. J., Squire, L. R., Paulsen, J. S., Swerdlow, N. R., Swenson, M., & Butters, N. (1996). Dissociations within non-declarative memory in Huntington's disease. *Neuropsychology*, *10*, 538-48.

Kobayashi, M., & Pascual-Leone, A. (2003). Transcranial magnetic stimulation in neurology. *Lancet Neurology*, *2*, 145-156.

Koch, G., Brusa, L., Carrillo, F., Lo Gerfo, E., Torriero, S., Oliveri, M., Mir, P., Caltagirone, C., & Stanzione, P. (2009). Cerebellar magnetic stimulation decreases levodopa-induced dyskinesias in Parkinson disease. *Neurology*, *73*, 113-119.

Koch, G., Mori, F., Marconi, B, Codecà, C., Pecchioli, C., Salerno, S., Torriero, S., Lo Gerfo, E., Mir, P., Oliveri, M., & Caltagirone, C. (2008). Changes in intracortical circuits of the human motor cortex following theta burst stimulation of the lateral cerebellum. *Clinical Neurophysiology*, *119*, 2559-2569.

Koch, G., Oliveri, M., Torriero, S., Salerno, S., Gerfo, E. L., & Caltagirone, C. (2007). Repetitive TMS of cerebellum interferes with millisecond time processing. *Experimental Brain Research*, *179*, 291-299.

Kortmann, B., & König, E. (1992). Categorial reanalysis: the case of deverbal prepositions. *Linguistics*, *30*, 671-97.

Kurylowicz, J. (1965). The evolution of grammatical categories. *Diogenes*, *51*, 55-71.

Kuteva, T. (2001). *Auxiliation: An enquiry into the nature of grammaticalization*. Oxford: Oxford University Press.

Lancaster, J. S., & Barsalou, L. W. (1997). Multiple organizations of events in memory. *Memory*, *5*, 569-599.

Lang, C. E., & Bastian, A. J. (2002). Cerebellar damage impairs automaticity of a recently practiced movement. *Journal of Neurophysiology*, *87*(3), 1336-1347.

Langacker, R. W. (1977). Syntactic reanalysis. In C. Li (Ed.), *Mechanisms of syntactic change* (pp. 57-139). Austin, Texas: University of Texas Press.

Langacker, R. W. (2000). A dynamic usage-based model. In M. Barlow & S. Kemmer (Eds.), *Usage-based models of language* (pp. 1-63). Stanford, CA: CSLI Publications.

Larsell, O. (1937). The cerebellum. A review and interpretation. *Archives of Neurology & Psychiatry (Chicago), 38*, 580-607.

Larsell, O., & Jansen, J. (1972). *The comparative anatomy and histology of the cerebellum, Vol 3 III: The human cerebellum, cerebellar connections and cerebellar cortex*. Minneapolis: University of Minnesota Press.

Leech, G. (1992). 100 million words of English: the British National Corpus. *Language Research, 28*(1), 1-13.

Lehmann, C. (1995 [1982]). *Thoughts on grammaticalization*. Munich: Lincom Europa. (originally published as akup 48, Institut für Sprachwissenschaft, Universität zu Köln).

Lehmann, C. (2002). New reflections on grammaticalization and lexicalization [Electronic version]. In I. Wischer & G. Diewald (Eds.), *New reflections on grammaticalization* (pp. 1-18). Amsterdam & Philadelphia: John Benjamins. Retrieved from http:// www. christianlehmann. eu/ publ/ New_reflections_on_grammaticalization_and_lexicalization.pdf

Lehmann, C. (2004). Theory and method in grammaticalization [Electronic version]. *Zeitschrift fur Germanistische Linguistik, 32*(2), 152-187. Retrieved from http://www.christianlehmann.eu/publ/Theory&method_in_grammaticalization.pdf

Leiner, H. C., Leiner, A. L., & Dow, R. S. (1986). Does the cerebellum contribute to mental skills? *Behavioral Neuroscience, 100*(4), 443-454.

Leiner, H. C., Leiner, A. L., & Dow, R. S. (1989). Reappraising the cerebellum: what does the hindbrain contribute to the forebrain? *Behavioral Neuroscience, 103*(5), 998-1008.

Leiner, H. C., Leiner, A. L., & Dow, R. S. (1991). The human cerebro-cerebellar system: Its computing, cognitive, and language skills. *Behavioural Brain Research, 24*, 113-128.

Leiner, H. C., Leiner, A. L., & Dow, R. S. (1995). The underestimated cerebellum. *Human Brain Mapping, 2*, 244-254.

Levitt, J. J., McCarley, R. W., Nestor, P. G., Petrescu, C., Donnino, R., Hirayasu, Y., Kikinis, R., Jolesz, F. A., & Shenton, M. E. (1999). Quantitative volumetric MRI study of the cerebellum and vermis in schizophrenia: clinical and cognitive correlates. *American Journal of Psychiatry, 156*, 1105-1107.

Lewis, P. A., & Miall, R. C. (2003). Distinct systems for automatic and cognitively controlled time measurement: evidence from neuroimaging. *Current Opinion in Neurobiology, 13*, 250-255.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). The perception of the speech code. *Psychological Review, 74*, 431-461.

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

Lightfoot, D. (1979). *Principles of diachronic syntax*. New York: Cambridge University Press.

Lightfoot, D. (1991). *How to set parametres: Arguments from language change*. Cambridge, Massachusetts: MIT Press.

Linebarger, M. C., Schwartz, M. F., & Saffran, E. M. (1983). Sensitivity to grammatical structure in so-called agrammatic aphasics. *Cognition, 13*, 361-392.

Lockwood, W. B. (1968). *Historical German syntax.* Oxford: Clarendon Press.

Lomber, S. G. (1999). The advantages and limitations of permanent or reversible deactivation techniques in the assessment of neural function. *Journal of Neuroscience Methods, 86*, 109-118.

Luxemburg, R. (1904 [1961]). Leninism or Marxism? In *The Russian Revolution and Leninism Or Marxism?* (pp. 81-108). Ann Arbor: The University of Michigan Press.

MacKay, W. A., & Murphy, J. T. (1979). Cerebellar modulation of reflex gain. *Progress in Neurobiology, 13*, 361-417.

MacLullich, A. M. J., Edmond, C. L., Ferguson, K. J., Wardlaw, J. M., Starr, J. M., Seckl, J. R., & Deary, I. J. (2004). Size of the neocerebellar vermis is associated with cognition in healthy elderly men. *Brain and Cognition, 56*, 344-348.

Maher, B. A. (1968, November). The shattered language of schizophrenia. *Psychology Today*, 30 ff.

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.

Mariën, P., Engelborghs, S., Pickut, B. A., & De Deyn, P. P. (2000). Aphasia following cerebellar damage: Fact or fallacy? *Journal of Neurolinguistics*, *13*, 145-171.

Martin, P., & Albers, M. (1995). Cerebellum and schizophrenia: a selective review. *Schizophrenia Bulletin*, *21*(2), 241-250.

Marvel, C. L., Schwartz, B. L., & Isaacs, K. L. (2004). Word production deficits in schizophrenia. *Brain and Language*, *89*(1), 182-191.

McCormick, D. A, & Thompson, R. F. (1984). Cerebellum: essential involvement in the classically conditioned eyelid response. *Science*, *223*, 296-299.

McEvoy, C. L., & Nelson, D. L. (1982). Category name and instance norms for 106 categories of various sizes. *American Journal of Psychology*, *95*(4), 581-634.

McRae, K., Ferretti, T. R., & Amyote, L. (1997). Thematic roles as verb-specific concepts. *Language and cognitive processes*, *12* (2/3), 137-176.

McRae, K., Hare, M., Elman, J. L., & Ferretti, T. (2005). A basis for generating expectancies for verbs from nouns. *Memory and Cognition*, *33*(7), 1174-1184.

McRae, K., Hare, M., Ferretti, T. R., & Elman, J. L. (2001). Activating verbs from typical agents, patients, instruments, and locations via event schemas. In D. Moore & K. Stenning (Eds.), *Proceedings of the twenty-third annual conference of the cognitive science society* (pp. 617-622). Mahwah, New Jersey: Erlbaum.

Meyer, B. U., Britton, T. C., Kloten, H., Steinmetz, H. & Benecke, R. (1991). Coil placement in magnetic brain stimulation related to skull and brain anatomy. *Electroencephalography and Clinical Neurophysiology*, *81*, 38-46.

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, *90*(2), 227-234.

Miall R. C. (2001). Cerebellum. In *Encyclopaedia of neurological science.* Academic Press. [Electronic version] Retrieved from http: // prism. bham. ac. Uk/ pdf_files/ Miall_2002 Encycl_Neurol_Sci.pdf

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

Miall, R. C. (2007). *Behavioural, imaging and TMS studies of the cerebellum in sensory motor control* [Powerpoint slides], Leeds, 2007. Retrieved from http://www.srr.org.uk/documents/about/SRR_Leeds_2007.pdf

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), 2733-2744.

Miall, R. C., & Jenkinson, E. W. (2005). Functional imaging of changes in cerebellar activity related to learning during a novel eye-hand tracking task. *Experimental Brain Research*, *166*, 170-183.

Miall, R. C., & King, D. (2008). State estimation in the cerebellum. *Cerebellum*, *7*, 572-576.

Miall, R. C., Weir, D. J., Wolpert, D. M., & Stein, J. F. (1993). Is the cerebellum a Smith Predictor? *Journal of Motor Behavior*, *25*, 203-216.

Miall, R. C., & Wolpert, D. M. (1996). Forward models for physiological motor control. *Neural Networks*, *9*, 1265-1279.

Middleton, F. A., & Strick, P. L. (1994). Anatomical evidence for cerebellar and basal ganglia involvement in higher cognitive function. *Science*, *266*, 458-461.

Middleton, F. A., & Strick, P. L. (1997). Cerebellar output channels. *International Review of Neurobiology*, *41*, 61-82.

Middleton, F. A., & Strick, P. L. (1998). Cerebellar output: motor and cognitive channels. *Trends in Cognitive Sciences*, *2*(9), 348-354.

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), 700-712.

Mishkin, M., Malamut, B., & Bachevalier, J. (1984). Memories and habits: Two neural systems. In G. Lynch, J. L. McGaugh, & N. M. Weinberger (Eds.), *Neurobiology of learning and memory* (pp. 65-77). New York: Guildford.

Mishkin, M., & Petri, H. L. (1984). Memories and habits: some implications for the analysis of learning and retention. In L. R. Squire, & N. Butters (Eds.), *The neuropsychology of memory* (pp. 287-96). New York: Guilford.

Molinari, M., Filippini, V., & Leggio, M. G. (2002). Neuronal plasticity of interrelated cerebellar and cortical networks. *Neuroscience*, *111*(4), 863-870.

Molinari, M., Leggio, M. G., & Silveri, M. (1997). Verbal fluency and agrammatism. In J. D. Schmahmann (Ed.), R. J. Bradley, R. A. Harris, & P. Jenner (Series Eds.), *The cerebellum and cognition. International Review of Neurobiology, 41* (pp. 325-339). San Diego: Academic Press.

Molinari, M., Restuccia, D., & Leggio, M. G. (2009). State Estimation, Response Prediction, and Cerebellar Sensory Processing for Behavioral Control. *Cerebellum*, *8*, 399-402.

Moretti, R., Bava, A., Torre, P., Antonello, R. M., & Cazzato, G. (2002). Reading errors in patients with cerebellar vermis lesions. *Journal of Neurology, 249*(4), 1432-1459.

Moritz, S., Mersmann, K., Kloss, M., Jacobsen, D., Andresen, B., Krausz, M., Pawlik, K., & Naber, D. (2001). Enhanced semantic priming in thought-disordered schizophrenic patients using a word pronunciation task, *Schizophrenia Research*, *48*(2), 301-305.

Moss, H. E., Ostrin, R. K., Tyler, L. K., & Marslen-Wilson, W. D. (1995). Accessing different types of lexical semantic information: evidence from priming. *Journal of experimental psychology: Learning, Memory, and Cognition*, *21*(4), 863-883.

Müller, F., & Dichgans, J. (1994). Dyscoordination of pinch and lift forces during grasp in patients with cerebellar lesions. *Experimental Brain Research*, *101*, 485-492.

Nagel, M., & Zangemeister, W. H. (2003). The effect of transcranial magnetic stimulation over the cerebellum on the synkinesis of coordinated eye and head movements. *Journal of Neurological Sciences*, *213*, 35-45.

Natsopoulos, D. (1985). A verbal illusion in two languages. *Journal of Psycholinguistic Research*, *14*, 385-397.

Nelson, M. E., & Paulin, M. G. (1995). Neural simulations of adaptive reafference suppression in the elasmobranch electrosensory system. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology, 177*(6), 723-736.

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), 139-142.

Newmeyer, F. J. (1998). *Language form and language function*. Cambridge, MA: MIT Press.

Newmeyer, F. J. (2001). Deconstructing grammaticalization. *Language Sciences, 23*, 187-229.

Nicolle, S. (1998), A relevance theory perspective on grammaticalization. *Cognitive Linguistics*, *9*(1), 1-35.

Nicolle, S. (2007). The grammaticalization of tense markers: a pragmatic reanalysis. In: L. de Saussure, J. Moeschler, & G. Puskas (Eds.), *Tense, mood and aspect: Theoretical and descriptive issues* (pp. 47-65). Amsterdam/New York: Rodopi.

Nicolson, R. I., Daum, I., Schugens, M. M., Fawcett, A. J., & Schulz, A. (2002). Abnormal eyeblink conditioning for dyslexic children. *Experimental Brain Research, 143*, 42-50.

Nicolson, R. I., & Fawcett, A. J. (1990). Automaticity: A new framework for dyslexia research. *Cognition, 35*, 159-182.

Nicolson R. I., Fawcett A. J., Berry, E. L., Jenkins, I. H., Dean, P., & Brooks, D. J. (1999). Association of abnormal cerebellar activation with motor learning difficulties in dyslexic adults. *Lancet, 353*, 1662-1667.

Nicolson, R. I., Fawcett, A. J., & Dean, P. (2001). Developmental dyslexia: The cerebellar deficit hypothesis. *Trends in Neurosciences, 24*, 508-511.

Niyogi, P., & Berwick, R. (1997). Evolutionary consequences of language learning. *Linguistics and Philosophy*, *20*, 697-719.

Noback, C. R., & Demarest, R. J. (1981). *The human nervous system: basic principles of neurobiology*. Lisbon: Graw-Hill.

Nowak, D. A., Hermsdorfer, J., Marquardt, C., & Fuchs, H. H. (2002). Grip and load force coupling during discrete vertical arm movements with a grasped object in cerebellar atrophy. *Experimental Brain Research*, *145*, 28-39.

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

Nyffeler, T., Cazzoli, D., Hess, C. W., & Müri, R. M. (2009). One Session of Repeated Parietal Theta Burst Stimulation Trains Induces Long-Lasting Improvement of Visual Neglect. *Stroke, 40,* 2791-2796.

Nyffeler, T., Cazzoli, D., Wurtz, P., Lüthi, M., von Wartburg, R., Chaves, S., Déruaz, A., Hess, C. W., & Müri, R. M. (2008). Neglect-like visual exploration behaviour after theta burst transcranial magnetic stimulation of the right posterior parietal cortex. *European Journal of Neuroscience, 27,* 1809 -1813.

Nyffeler, T., Wurtz, P., Lüscher, H.R., Hess, C. W., Senn, W., Pflugshaupt, T., von Wartburg, R., Luthi, M., & Müri, R. M. (2006). Extending lifetime of plastic changes in the human brain. *European Journal of Neuroscience, 24,* 2961-6.

Ober, B. (2002). RT and non-RT methodology for semantic priming research with Alzheimer's disease patients: A critical review. *Journal of Clinical and Experimental Neuropsychology, 24*(7), 883-911.

Ohtsuka, K., & Enoki, T. (1998). Transcranial magnetic stimulation over the posterior cerebellum during smooth pursuit eye movements in man. *Brain, 121,* 429-435.

Ohyama, T., Nores, W. L., Murphy, M., & Mauk, M. D. (2003). What the cerebellum computes. *Trends in Neurosciences, 26*(4), 222-227.

Okugawa, G., Nobuhara, K., Takase, K., & Kinoshita, T. (2007). Cerebellar posterior superior vermis and cognitive cluster scores in drug-naïve patients with first-episode schizophrenia. *Neuropsychobiology, 56,* 216-219.

Okugawa, G., Sedvall, G. C., & Agartz, I. (2003). Smaller cerebellar vermis but not hemisphere volumes in patients with chronic schizophrenia. *American Journal of Psychiatry, 160,* 1614-1617.

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., Torriero, S., Koch, G., Salerno, S., Petrosini, L., & Caltagirone, C. (2007). The role of transcranial magnetic stimulation in the study of cerebellar cognitive function. *The Cerebellum, 6,* 95-101.

O' Shea, J., & Walsh, V. (2007). Transcranial magnetic stimulation. *Current Biology, 17*(6), 196-199.

Palmer, E. D., Rosen, H. J., Ojemann, J. G., Buckner, R. L., Kelley, W. M., & Petersen, S. E. (2001). An event-related fMRI study of overt and covert word stem completion. *NeuroImage, 14,* 182-193.

Pascual-Leone, A., Bartres-Faz, D., & Keenan, J. P. (1999).Transcranial magnetic stimulation: studying the brain-behaviour relationship by induction of 'virtual lesions'. *Philosophical Transactions of the Royal Society of London.B, 354,* 1229-1238.

Pascual-Leone, A., Houser, C. M., Reese, K., Shotland, L. I., Grafman, J., Sato, S., Valls-Sole, J., Brasil-Neto, J. P., Wasserman, E. M., Cohen, L. G., & Hallett, M. (1993). Safety of rapid-rate transcranial magnetic stimulation in normal volunteers. *Electroencephalography and Clinical Neurophysiology, 89,* 120-130.

Pascual-Leone, A., Tormos, J.M., Keenan, J., Tarazona F., Canete C., & Catala, M.D. (1998). Study and modulation of human cortical excitability with transcranial magnetic stimulation. *Journal of Clinical Neurophysiology, 15,* 333-343.

Pascual-Leone, A., Walsh, V., & Rothwell, J. (2000). Transcranial magnetic stimulation in cognitive neuroscience- virtual lesion, chronometry, and functional connectivity. *Current Opinion in Neurobiology, 10,* 232-237.

Passingham, R. E. (1975). Changes in the size and organization of the brain in man and his ancestors. *Brain Behavioral Evolution, 11,* 73-90.

Paulin, M. G. (1989). A Kalman filter theory of the cerebellum. In M. A. Arbib & S.- I. Amari (Eds.), *Dynamic interactions in neural Networks: Models and data* (pp. 239-259). New York: Springer.

Paulin, M. G. (1993). The role of the cerebellum in motor control and perception. *Brain, Behavior, and Evolution. 41,* 39-50.

Paulin, M. G. (1996). Cerebellar theory out of control. *Behavioral and Brain Sciences. 19,* 470-471.

Paulin, M. G. (1997). Neural representations of moving systems. In J.D. Schmahmann (Ed.), R. J. Bradley, R. A. Harris, & P. Jenner (Series Eds.), *The cerebellum and cognition. International Review of Neurobiology, 41* (pp. 515-533). San Diego: Academic Press.

Paus, T. (1999). Imaging the brain before, during, and after transcranial magnetic stimulation. *Neuropsychologia, 37,* 219-224.

Pedersen, T., Patwardhan, S., & Michelizzi, J. (2004). Wordnet::similarity- measuring the relatedness of concepts. *Proceedings of the nineteenth national conference on artificial intelligence (AAAI-04),* 1024-1025.

Peirce, C. S. (1931-58). *Collected Writings* (Vol. 1-8). C. Hartshorne, P. Weiss, & A. W. Burks (Eds.). Cambridge, MA: Harvard University Press.

Perea, M., & Gotor, A. (1997). Associative and semantic priming effects occur at very short stimulus-onset asynchronies in lexical decision and naming. *Cognition*, *62*, 223-240.

Perea, M., & Rosa, E. (2002). The effects of associative and semantic priming in the lexical decision task. *Psychological Research, 66,* 180-194.

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.

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*, 153-170.

Petersson, K. M., Folia, V., & Hagoort, P. (2010). *What artificial grammar learning reveals about the neurobiology of syntax.* Manuscript submitted for publication.

Petersson, K. M., Forkstam, C., & Ingvar, M. (2004). Artificial syntactic violations activate Broca's area. *Cognitive Science*, *28*, 383-407.

Petrides, M., & Pandya, D. N. (1994). Comparative architectonic analysis of the human and macaque frontal cortex. In F. Boller & J. Grafman (Eds.), *Handbook of neuropsychology, Vol. 9* (pp. 17-58). Amsterdam: Elsevier.

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), 105-110.

Pickering, M. J., & Garrod, S. (2009a). Prediction and embodiment in dialogue. *European Journal of Social Psychology*, *39*, 1162-1168.

Pickering, M. J., & Garrod, S. (2009b). Language, interaction, and embodiment. *European Journal of Social Psychology*, *39*, 1178-1179.

Pickett, E. R. (1998). *Language and the cerebellum*. Unpublished doctoral dissertation, Brown University, Providence, Rhode Island.

Poldrack, R. A., Prabhakaran, V., Seger, C. A., & Gabrieli, J. D. (1999). Striatal activation during acquisition of a cognitive skill. *Neuropsychology, 13(4),* 564-574.

Pollatsek, A., & Well, A. D. (1995). On the use of counterbalanced designs in cognitive research: A suggestion for a better and more powerful analysis. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 21*, 785-794.

Pothos, E. M., & Chater, N. (2002). A simplicity principle in unsupervised human categorization. *Cognitive Science, 26*, 303–343.

Psychology Software Tools, Inc. (2009). *E-Prime* (version 2.0) [Computer software] [http://www.pst-net.com/]. Pittsburgh, PA, USA.

Pulvermüller, F. (2002). *The neuroscience of language: On brain circuits of words and serial order*. Cambridge: Cambridge University Press.

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

Ramnani, N., & Miall, R. C. (2001). Expanding cerebellar horizons. *Trends in Cognitive Sciences, 5*(4), 135-136.

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior, 5*, 855-863.

Richter, S., Kaiser, O., Hein-Kropp, C., Dimitrova, A., Gizewski, E., Beck, A., Aurich, V., Ziegler, W., & Timmann, D. (2004). Preserved verb generation in patients with cerebellar atrophy. *Neuropsychologia, 42*, 1235-1246.

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.

Rizzolatti, G. (2008). *Mirror neurons: a paradigm shift in cognitive science?* Presentation in first interdisciplinary PPLS Seminar, 2 September 2008. University of Edinburgh.

Roberts, I. (1993). A formal account of grammaticalization in the history of Romance futures. *Folia Linguistica Historica, 13(1-2),* 219-258.

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

Rosenbach, A., & Jäger, G. (2008). Priming and unidirectional language change. *Theoretical Linguistics, 34*(2), 85-113.

Rossini, P. M., Barker, A. T., Berardelli, A., Caramia, M. D., Caruso, G., Cracco, R. Q., Dimitrijević, M. R., Hallett, M., Katayama, Y., Lücking, C. H., Maertens de Noordhout, A. L., Marsden, C. D., Murray, N. M. F., Rothwell, J. C., Swash, M., & Tomberg, C. (1994). Noninvasive electrical and magnetic stimulation of the brain, spinal cord and roots: basic principles and procedures for routine clinical application. Report of an IFCN committee. *Electroencephalography and Clinical Neurophysiology, 91*, 79-92.

Roth, B. J., Saypol, J. M., Hallett, M., & Cohen, L. G. (1991). A theoretical calculation of the electric field induced in the cortex during magnetic stimulation. *Electroencephalography and Clinical Neurophysiology, 81*, 47-56.

Rothwell, J. C. (1997). Techniques and mechanisms of action of transcranial stimulation of the human motor cortex. *Journal of Neuroscience Methods, 74*, 113-122.

Rudiak, D., & Marg, E. (1994). Finding the depth of magnetic brain stimulation: a re-evaluation. *Electroencephalography and Clinical Neurophysiology, 93*, 358-371.

Rumelhart, D. E. (1980). Schemata: The building blocks of cognition. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension* (pp. 33-58). Hillsdale, New Jersey: Erlbaum.

Rumelhart, D. E., & Norman, D. A. (1982). Simulating a skilled typist: A study of skilled cognitive-motor performance. *Cognitive Science, 6*, 1-36.

Sack, A. T. (2006). Transcranial magnetic stimulation, causal structure-function mapping and networks of functional relevance. *Current Opinion in Neurobiology, 16*, 593-599.

Saffran, E. M., Schwartz, M. F., & Linebarger, M. C. (1998). Semantic influences on thematic role assignment: Evidence from normals and aphasics. *Brain and Language, 62*, 255-297.

Samuel, A. G. (1981). Phonemic restoration: Insights from a new methodology. *Journal of Experimental Psychology: General, 110*, 474-494.

Samuel, A. G., & Troicki, M. (1998). Articulation quality is inversely related to redundancy when children or adults have verbal control. *Journal of memory and language, 39*, 175-194.

Sanford, A. J., & Graesser, A. (2006). Introduction: Shallow processing and underspecification. *Discourse Processes, 42*(2), 99-108.

Sanford, A. J., & Sturt, P. (2002). Depth of processing in language comprehension: not noticing the evidence. *Trends in Cognitive Sciences, 6*(9), 382-386.

Schindler, K., Nyffeler, T., Wiest, R., Hauf, M., Mathis, J., Hess, Ch. W., & Müri, R. (2008). Theta burst transcranial magnetic stimulation is associated with increased EEG synchronization in the stimulated relative to unstimulated cerebral hemisphere. *Neuroscience Letters, 436*, 31-34.

Schmahmann, J. D. (1996). From movement to thought: Anatomic substrates of the cerebellar contribution to cognitive processing. *Human Brain Mapping, 4*, 174-198.

Schmahmann, J. D. (1998). Dysmetria of thought. Clinical consequences of cerebellar dysfunction in cognition and affect. *Trends in Cognitive Sciences, 2*, 362-370.

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

Schmahmann, J. D. (2001). The cerebrocerebellar system: anatomic substrates of the cerebellar contribution to cognition and emotion. *International Review of Psychiatry, 13*, 247-260.

Schmahmann, J. D., & Caplan, D. (2006). Cognition, emotion and the cerebellum. *Brain, 129*(2), 290-292.

Schmahmann, J. D., & Pandya, D. N. (1989). Anatomical investigation of projections to the basis pontis from posterior parietal association cortices in rhesus monkey. *Journal of Comparative Neurology, 289*(1), 53-73.

Schmahmann, J. D., & Pandya, D. N. (1995). Prefrontal cortex projections to the basilar pons in rhesus monkey: implications for the cerebellar contribution to higher function. *Neuroscience Letters, 199*(3), 175-178.

Schmahmann, J. D., & Pandya, D. N. (1997). The cerebrocerebellar system. In J. D. Schmahmann (Ed.), R. J. Bradley, R. A. Harris, & P. Jenner (Series Eds.), *The cerebellum and cognition. International Review of Neurobiology, 41* (pp. 31-60). San Diego: Academic Press.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, *84*, 1-66.

Shadmehr, R., & Holcomb, H. H. (1997). Neural correlates of motor memory consolidation. *Science, 277*, 821-825.

Shah, V. S., Schmahmann, J. D., Pandya, D. N., & Vaher, P. R. (1997). Associative projections to the zona incerta: Possible anatomic substrates for extension of the Marr-Albus hypothesis to non-motor learning. *Society of Neuroscience Abstracts*, *23*, 1829.

Sherwood, L. (2010). *Human physiology: From cells to systems* (7th ed.). Pacific Grove, California: Brooks/Cole.

Silvanto, J., Muggleton, N. G., Cowey, A., & Walsh, V. (2007). Neural activation state determines behavioral susceptibility to transcranial magnetic stimulation. *European Journal of Neuroscience*, *26*, 523-528.

Silveri, M. C., Leggio, M. G., & Molinari, M. (1994). The cerebellum contributes to linguistic production: A case of agrammatic speech following a right cerebellar lesion. *Neurology*, *44*, 2047-2050.

Silveri, M. C., & Misciagna, S. (2000). Language, memory and cerebellum. *Journal of Neurolinguistics*, *13*, 129-143.

Simon, D. (2006). *Optimal state estimation: Kalman, H∞, and nonlinear approaches*. Hoboken, New Jersey: Wiley.

Smith, K., Kirby, S., & Brighton, H. (2003). Iterated learning: A framework for the emergence of language. *Artificial Life, 9(4)*, 371-386.

Sperling, A. J., Lu, Z.- L., Manis, F. R., & Seidenberg, M. S. (2005). Deficits in perceptual noise exclusion in developmental dyslexia. *Nature Neuroscience*, *8*(7), 862-863.

Sternberg, S. (1966). High-speed scanning in human memory. *Science, 153*, 652- 654.

Stewart, L., Walsh, V., Frith, U., & Rothwell, J. C. (2001). TMS produces two dissociable types of speech disruption. *NeuroImage*, *13*, 472-478.

Stowe, L. A., Paans, A. M. J., Wijers, A. A., & Zwarts, F. (2004). Activations of "motor" and other non-language structures during sentence comprehension. *Brain and Language*, *89*, 290-299.

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

Thach, W. T. (1996a). On the specific role of the cerebellum in motor learning and cognition: Clues from PET activation and lesion studies in man. *Behavioral and Brain Sciences*, *19*(3), 411-431.

Thach, W.T. (1996b). Q. Is the cerebellum an adaptive combiner of motor and mental/motor activities? A. yes, maybe, certainly not, who can say? *Behavioral and Brain Sciences*, *19*, 501-503.

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

Thach, W. T. (1998). What is the role of the cerebellum in motor learning and cognition? *Trends in Cognitive Sciences*, *2*(9), 331-337.

Thach, W.T. (2007). On the mechanism of cerebellar contributions to cognition. *The Cerebellum*, *6*, 163-167.

Thach, W. T., Goodkin, H. P., & Keating, J. G. (1992). The cerebellum and the adaptive coordination of movement. *Annual Review of Neuroscience*, *15*, 403-442.

Thach, W. T., Kane, S. A., & Goodkin, H. P. (1998). Dentate/interposed nuclear sites coordinating reach and grasp. *Society of Neuroscience Abstracts*, *24*, 1406.

Thach, W. T., Mink, J. W., Goodkin, H. P., & Keating J. G. (1993). Combining versus gating motor programs: differential roles for cerebellum and basal ganglia? In M. Mano, I. Hamada, & M. R. DeLong (Eds.), *Roles of cerebellum and basal ganglia in voluntary movement* (pp. 235-245). Amsterdam: Elsevier.

The Magstim Company Ltd (2006). *Magstim® Rapid² P/N 3576-23-08 Operating Manual.* Retrieved from http://www.magstim.com/15114.file.dld

The Magstim Company Ltd (2009). *Website, frequently asked questions.* Retrieved from http://www.magstim-us.com/support/mfaq8.html

Théoret, H., Haque, J., & Pascual-Leone, A. (2001). Increased variability of paced finger tapping accuracy following repetitive magnetic stimulation of the cerebellum in humans. *Neuroscience Letters*, *306*, 29-32.

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. In J. D. Schmahmann (Ed.), R. J. Bradley, R. A. Harris, & P. Jenner (Series Eds.),

The cerebellum and cognition. *International Review of Neurobiology*, *41* (pp. 152-189). San Diego: Academic Press.

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*, 440-458.

Timmann, D., & Daum, I. (2007). Cerebellar contributions to cognitive functions: A progress report after two decades of research. *The Cerebellum*, *6*, 159-162.

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*, 2056-2060.

Todd, G., Flavel, S. C., & Ridding, M. C. (2009). Priming theta-burst repetitive transcranial magnetic stimulation with low- and high-frequency stimulation. *Experimental Brain Research*, *195*(2), 307-315.

Tomasello, M. (2003). On the different origins of symbols and grammar. In M. H. Christiansen & S. Kirby (Eds.), *Language evolution* (pp. 94-110). Oxford: Oxford University Press.

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.

Torriero, S., Oliveri, M., Koch, G., Caltagirone, C., & Petrosini, L. (2004). Interference of left and right cerebellar rTMS with procedural learning. *Journal of Cognitive Neuroscience*, *16*, 1605-1611.

Torriero, S., Oliveri, M., Koch, G., Lo Gerfo, E., Salerno, S., Petrosini, L., & Caltagirone, C. (2007). Cortical networks of procedural learning: evidence from cerebellar damage. *Neuropsychologia*, *45*(6), 1208-1214.

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

Traugott, E. C., & Dasher, R. B. (2005). *Regularity in semantic change*. Cambridge: Cambridge University Press.

Traugott, E. C, & König, E. (1991). The semantics-pragmatics of grammaticalization revisited. In E. C. Traugott & B. Heine (Eds.), *Approaches to grammaticalization* (pp. 189-218). Amsterdam: John Benjamins.

Tsangalidis, A. (1999). *Will and tha: a comparative study of the category future*. Thessaloniki: University Studio 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*, 1229-1234.

Turk A. (2010). *Does prosodic constituency reflect predictability?* Manuscript submitted for publication.

Turton, D., & Bender, M. L. (1976). Mursi. In M. L. Bender (Ed.), *The non-Semitic languages of Ethiopia* (pp. 533-61). East Lansing, MI: African Studies Center, Michigan State University Press.

Udden, J., Folia, V., Forkstam, C., Ingvar, M., Fernandez, G., Overeem, S., van Elswijk, G., Hagoort, P., & Petersson, K. M. (2008). The inferior frontal cortex in artificial syntax processing: An rTMS study. *Brain Research*, *1224*, 69-78.

Ugawa, Y., Day, B. L., Rothwell, J. C., Thompson, P. D., Merton, P. A., & Marsden, C. D. (1991). Modulation of motor cortical excitability by electrical stimulation over the cerebellum in man. *Journal of Physiology (London)*, *441*, 57-72.

Ugawa, Y., Uesaka, Y., Terao Y., Hanajima, R., & Kanazawa, I. (1995). Magnetic stimulation over the cerebellum in humans. *Annals of Neurology*, *37*(6), 703-713.

Ullman, M. T. (2004). Contributions of memory circuits to language: the declarative/procedural model. *Cognition*, *92*, 231-270.

Vallesi, A., Shallice, T., & Walsh, V. (2007). Role of the prefrontal cortex in the foreperiod effect: TMS evidence for dual mechanisms in temporal preparation. *Cerebral Cortex*, *17*, 466-474.

Van Overschelde, J. P., Rawson, K. A., & Dunlosky, J. (2004). Category norms: An updated and expanded version of the Battig and Montague (1969) norms. *Journal of Memory and Language*, *50*, 289-335.

Vega Moreno, R. (2007). *Creativity and convention*. Amsterdam: John Benjamins.

Voss, M., Bays, P. M., Rothwell, J. C., & Wolpert, D. M. (2008). An improvement in perception of self-generated tactile stimuli following theta-burst stimulation of primary motor cortex. *Neuropsychologia*, *45*(12), 2712-2717.

Vosse, T., & Kempen, G. (2000). Syntactic structure assembly in human parsing: a computational model based on competitive inhibition and a lexicalist grammar. *Cognition*, *75*, 105-143.

Walenski, M., Mostofsky, S. H., & Ullman, M. T. (2007). Speeded processing of grammar and tool knowledge in Tourette's syndrome. *Neuropsychologia, 45*, 2447-2460.

Walsh, V., & Cowey, A. (2000). Transcranial magnetic stimulation and cognitive neuroscience. *Nature Reviews Neuroscience, 1*, 73-79.

Walsh, V., & Pascual-Leone, A. (2002). Case studies in virtual neuropsychology: reversible lesions and magnetic brain stimulation. In R. Galuske & S. Lomber (Eds.), *Reversible lesion methods* (pp. 249-284). Oxford: Oxford University Press.

Walsh, V., & Pascual-Leone, A. (2003). *Transcranial magnetic stimulation: A neurochronometrics of mind.* Boston, Massachusetts: MIT Press.

Walsh, V., & Rushworth, M. (1999). A primer of magnetic stimulation as a tool for neuropsychology. *Neuropsychologia*, *37*, 125-135.

Ward, P. (2007). *The time course of sentence interpretation*. Unpublished doctoral dissertation, University of Glasgow, Scotland.

Warren, R. M. (1970). Perceptual restoration of missing speech sounds. *Science, 167*, 392-393.

Wason, P., & Reich, S. S. (1979). A verbal illusion. *Quarterly Journal of Experimental Psychology, 31*, 591-597.

Wassermann, E. M. (1998). Risk and safety of repetitive transcranial magnetic stimulation: Report and suggested guidelines from the International Workshop on the Safety of Repetitive Transcranial Magnetic Stimulation, June 5-7, 1996. *Electroencephalography and Clinical Neurophysiology*, *108*, 1-16.

Webb, B. (2004). Neural mechanisms for prediction: do insects have forward models? *Trends in Neurosciences, 27*(5), 278-282.

Werhahn, K. J., Taylor, J., Ridding, M., Meyer, B. U., & Rothwell, J. C. (1996). Effect of transcranial magnetic stimulation over the cerebellum on the excitability of human motor cortex. *Electroencephalography and Clinical Neurophysiology*, *101*, 58-66.

Williams, L. E., Ramachandran, V. S., Hubbard, E. M., Braff, D. L., & Light, G. A. (2010). Superior size-weight illusion performance in patients with schizophrenia: Evidence for deficits in forward models. *Schizophrenia Research*, *121*, 101-106.

Williams, R. W., & Herrup, K. (1988). The control of neuron number. *Annual Review of Neuroscience*, *11*, 423-453.

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

Witt, K., Nushman, A., & Deuschl, G. (2002). Intact artificial grammar learning in patients with cerebellar degeneration and advanced Parkinson's disease. *Neuropsychologia*, *40*(9), 1534-1540.

Wolpert, D. M. (1997). Computational approaches to motor control. *Trends in Cognitive Sciences*, *1*, 209-216.

Wolpert, D. M., Miall, R. C., & Kawato, M. (1998). Internal models in the cerebellum. *Trends in Cognitive Sciences*, *2*(9), 338-347.

Woodruff-Pak, D. S., Papka, M., & Ivry, R. B. (1996). Cerebellar involvement in eyeblink classical conditioning in humans. *Neuropsychology*, *10*, 443-458.

Woodruff-Pak, D. S., Vogel III, R. W., Ewers, M., Coffey, J., Boyko, O. B., & Lemieux, S. K. (2001). MRI-assessed volume of cerebellum correlates with associative learning. *Neurobiology of Learning and Memory*, *76*, 342-357.

Wray, A., & Perkins, M. R. (2000). The functions of formulaic language: An integrated model. *Language and Communication*, *20*, 1-28.

Xiang, H., Lin, C., Ma, X., Zhang, Z., Bower, J. M., Weng, X., & Gao, J. H. (2003). Involvement of the cerebellum in semantic discrimination: an fMRI study. *Human Brain Mapping*, *18*, 208-214.

Zagon, I. S., MacLaughlin, P. J., & Smith, S. (1977). Neuronal population in the human cerebellum: Estimations from isolated cell nuclei. *Brain Research*, *127*, 279-282.

Zangen, A., Roth, Y., Voller, B., & Hallett, M. (2005). Transcranial magnetic stimulation of deep brain regions: evidence for efficacy of the H-Coil. *Clinical Neurophysiology*, *116*, 775-779.

Ziegler, J. C., Pech-Georgel, C., George, F., & Lorenzi, C. (2009). Speech-perception-in-noise deficits in dyslexia. *Developmental Science*, *12*(5), 732-745.

Zuraw, K. (2003). Probability in language change. In R. Bod, J. Hay, & S. Jannedy (Eds.), *Probabilistic linguistics* (pp. 139-176). Cambridge, Massachusetts: MIT Press.

Appendices

Appendix A

Material for Experiments 1, 2, and 3

(TMS study 1)

Formal-Associatively Related Pairs		Semantic-Categorically Related Pairs	
Prime	**Target**	**Prime**	**Target**
tummy	ACHE	orange	CITRUS
estate	AGENT	pants	CLOTHING
snake	BITE	penny	COIN
sand	CASTLE	knife	CUTLERY
trench	COAT	cold	DISEASE
prawn	COCKTAIL	kettle	DRUM
battle	CRY	apartment	DWELLING
ostrich	FEATHER	cotton	FABRIC
passion	FLOWER	toilet	FIXTURE
blast	FURNACE	sandals	FOOTWEAR
pigeon	HOLE	mountain	FORMATION
cape	HORN	apple	FRUIT
gift	HORSE	diamond	GEM
scrap	IRON	mint	HERB
traffic	JAM	wasp	INSECT
tomato	JUICE	teller	JOB
swan	LAKE	whisky	LIQUOR
cabinet	MAKER	breakfast	MEAL
test	MATCH	telephone	MEDIUM
bridle	PATH	governor	POLITICIAN
rice	PUDDING	dictionary	REFERENCE
protection	RACKET	crab	SEAFOOD
cash	REGISTER	yacht	SHIP
slate	ROOF	football	SPORT
foam	RUBBER	ball	TOY
computer	SCIENCE	mahogany	TREE
microscope	SLIDE	foot	UNIT
lawn	TENNIS	bike	VEHICLE
shock	TREATMENT	canal	WATERWAY
maternity	WARD	maple	WOOD

Table A.1 : Unrelated and related pairs for formal-associative and semantic-categorical sets (experiments 1, 2, and 3).

Blocks	Prime				Target			
	Frequency (words/million)		Length (number of letters)		Frequency (words/million)		Length (number of letters)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
1	33.59	45.48	5.73	1.64	37.87	32.12	5.58	1.68
2	32.04	37.96	5.73	1.52	38.55	51.24	5.58	1.73

Table A.2: Length and Frequency for all pairs across the two blocks (experiments 1, 2, and 3).

Pairs	Related								Unrelated							
BLOCKS	Prime				Target				Prime				Target			
	Frequency (w/million)		Length (number of letters)		Frequency (w/million)		Length (number of letters)		Frequency (w/million)		Length (number of letters)		Frequency (w/million)		Length (number of letters)	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
1	35.7	38.6	5.6	1.6	37.8	30.4	5.2	1.6	34.1	50.4	5.5	1.5	37.8	30.4	5.2	1.6
2	33.8	52.1	5.8	1.7	37.1	38	5.2	1.2	31.5	37.2	5.6	1.5	37.1	38.0	5.2	1.2

Table A.3: Word length and frequency for related and unrelated pairs of the formal-associative group (experiments 1, 2, and 3).

Pairs	Related								Unrelated							
BLOCKS	Prime				Target				Prime				Target			
	Frequency (words/million)		Length (number of letters)		Frequency (words/million)		Length (number of letters)		Frequency (words/million)		Length (number of letters)		Frequency (words/million)		Length (number of letters)	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
1	30.9	34.3	5.83	1.9	37.9	34.5	6.0	1.7	33.7	57.2	5.9	1.5	37.9	34.5	6.0	1.7
2	31.3	27.6	5.9	1.5	40.0	62.8	5.9	2.0	31.6	32.3	5.7	1.5	40.0	62.8	6.0	2.1

Table A.4: Word length and frequency for related and unrelated pairs of the semantic-categorical group (experiments 1, 2, and 3).

Priming	Block	EAT		Wordnet::Similarity		BNC (words/million)	
		Mean	SD	Mean	SD	Mean	SD
Formal-Associative	1	0.17	0.07	0.12	0.03	0.55	0.71
	2	0.16	0.05	0.12	0.03	0.57	0.89
Semantic-Categorical	1	0.01	0.02	0.42	0.09	0.01	0.03
	2	0.01	0.02	0.41	0.08	0.00	0.01

Table A.5: Semantic similarity, lexical co-occurrence and word association strength for formal-associatively and semantic-categorically related pairs (experiments 1, 2, and 3; see section 5.2.1 for dependent measures).

Analysis of latencies for the first TMS study

Mean latencies per TMS Site (NCBV, PLCB), Priming Type (formal-associative, semantic-categorical), TMS Phase (before, after TMS), and Relatedness (unrelated, related) are demonstrated in table A.6 and plot A.1 below.

A **4-way ANOVA (site, priming type, phase, relatedness)** demonstrated an expected acceleration of reaction times in the second phase, marginally by subjects and significantly by items (phase: F_1 (1, 6) = 5.68, MSe = 3433.56, p = .06; F_2 (1, 112) = 36.98, MSe = 2040.85, p < .001). Latencies for the associative group were much shorter than those for the semantic group, significantly so by subjects (priming type: F_1 (1, 6) = 9.68, MSe = 540.69, p < .05; F_2 (1, 112) = 2.51, MSe = 13170.10, p = .12). Sadly, also, the number of participants here, unlike the larger pilot group in the first study, did not suffice to show a relatedness effect across priming types (F_1 (1, 6) = 2.81, MSe = 439.89, p = .15; F_2 (1, 112) = 1.4, MSe = 7478.32, p > .2). The analysis also showed a significant by-items site*phase interaction (F_1 (1, 6) = 1.4, MSe = 2067.60, p > .3; F_2 (1, 112) = 7.53, MSe = 1593.74, p < .01). This is because the acceleration of responses after PLCB TMS is much stronger than that after NCBV TMS (see below). Crucially, though, there was also **a strong site*priming type*phase*relatedness interaction (F_1 (1, 6) = 20.25, MSe = 165.55, p < .005; F_2 (1, 112) = 9.27, MSe = 1593.74, p < .005**; no site*phase*relatedness interaction: F_1 (1, 6) = 2.25, MSe = 95.55, p = .18; F_2 < 1; rest of ps, p > .2). This was at least partly owed to the **selective decrease of RTs for associatively related items after NCBV TMS**.

3-way ANOVAs: A comparison of **NCBV TMS** conditions **(priming type, phase, relatedness)** showed **a strong priming type*phase*relatedness interaction (F_1 (1, 6) = 14.31, MSe = 190.95, p < .01; F_2 (1, 112) = 7.45, MSe = 1759.54, p < .01**; also, phase: F_1 (1, 6) = 2.2, MSe = 2094.74, p > .2. F_2 (1, 112) = 7.75, MSe = 1759.54, p < .01; priming type: F_1 (1, 6) = 7.88, MSe = 435.94, p < .05, F_2 (1, 112) = 1.8, MSe =

5568.44, $p > .2$; relatedness: F_1 (1, 6) = 5.17, MSe = 267.31, $p = .06$; F_2 (1, 112) = 1.73, $p = .19$; priming type*relatedness: $F_1 < 1$; F_2 (1, 112) = 1.80, MSe = 6262.42, $p = .18$; rest of *F*s, $F < 1$).

3-way ANOVAs: On the contrary, a comparison of RTs for **PLCB TMS** conditions **(priming type, phase, relatedness)** demonstrated an only marginal by subjects phase*priming type*relatedness interaction (F_1 (1, 6) = 4.62, MSe = 189.88, $p = .08$; F_2 (1, 112) = 1.76, MSe = 1875.05, $p = .19$; also, phase: F_1 (1, 6) = 5.17, MSe = 3406.42, $p = .06$; F_2 (1, 112) = 39.38, MSe = 1875.05, $p < .001$).

3-way ANOVAs: A comparison of RTs for the phase **after TMS (site, priming type, relatedness)** showed **a clear site*priming type*relatedness interaction (F_1 (1, 6) = 8.13, MSe = 227.18, $p < .05$; F_2 (1, 112) = 8.91, MSe = 1626.22, $p < .005$**; also, site: $F_1 < 1$; F_2 (1, 112) = 2.68, MSe = 1626.22, $p = .1$; site*relatedness: F_1 (1, 6) = 8.66, MSe = 59.37, $p < .05$; $F_2 < 1$; priming type*relatedness: F_1 (1, 6) = 2.78, MSe = 193.50, $p = .15$; F_2 (1, 112) = 1.4, MSe = 5363.30, $p > .25$; rest of *p*s, $p > .2$).

3-way ANOVAs: On the contrary, a comparison of RTs for the phase **before TMS (site, priming type, relatedness)** demonstrated an only marginal by subjects site*priming type*relatedness interaction (F_1 (1, 6) = 4.34, MSe = 348.91, $p = .08$; F_2 (1, 112) = 1.91, MSe = 1392.44, $p = .17$; also, site: $F_1 < 1$; F_2 (1, 112) = 5.68, MSe = 1392.44, $p < .05$; relatedness: both *F*s, $p > .3$; rest of *F*s, $F < 1$).

3-way ANOVAs: A comparison of RTs for **formal-associative priming** conditions **(site, phase, relatedness)** showed **a significant site*phase*relatedness interaction (F_1 (1, 6) = 19.10, MSe = 137.86, $p < .01$; F_2 (1, 56) = 4.51, MSe = 1825.24, $p < .05$**; also, site: $F_1 < 1$; F_2 (1, 56) = 25.64, MSe = 1853.37, $p < .001$; phase: F_1 (1, 6) = 6.45, MSe = 1701.35, $p < .05$; $F_2 < 1$; relatedness: F_1 (1, 6) = 7.41, MSe = 182.21, $p < .05$; F_2 (1, 56) = 2.74, MSe = 4097.77, $p = .1$; rest of *p*s, $p > .2$).

3-way ANOVAs: A comparison of RTs for **semantic-categorical priming** conditions **(site, phase, relatedness)** showed an also significant site*phase*relatedness interaction (F_1 (1, 6) = 7.58, MSe = 123.24, $p < .05$; F_2 (1, 56) = 4.83, MSe = 1362.24, $p < .05$; also, phase: F_1 (1, 6) = 3.90, MSe = 2206.09, $p = .1$; F_2 (1, 56) = 13.05, MSe = 2228.33, $p < .005$; site*phase: F_1 (1, 6) = 1.4, MSe = 1253.71, $p > .25$; F_2 (1, 56) = 4.61, MSe = 1362.24, $p < .05$; rest of *F*s, $F < 1$).

3-way ANOVAs: A comparison of RTs **for related items (site, priming type, phase)** demonstrated **a site*priming type*phase interaction, which was significant by items, but marginal by subjects (F_1 (1, 6) = 4.31, MSe = 558.44, $p = .08$; F_2 (1, 56) = 5.86, MSe = 1311.68, $p < .05$)**, and a marginal by items site*phase interaction ($F_1 < 1$; F_2 (1, 56) = 4.02, MSe = 1311.68, $p = .05$; also, phase: F_1 (1, 6) = 3.62, MSe = 2788.85, $p = .11$; F_2 (1, 56) = 24.03, MSe = 1550.05, $p < .001$; no phase*priming type interaction: $F_1 < 1$; F_2 (1, 56) = 1.76, MSe = 1550.05, $p = .19$; rest of *F*s, $F < 1$).

3-way ANOVAs: A comparison of RTs for **unrelated items (site, priming type, phase)** also demonstrated a site*priming type*phase interaction, significantly by subjects, and marginally by items (F_1 (1, 6) = 6.75, MSe = 159.83, $p < .05$; F_2 (1, 56) = 3.78, MSe = 1875.80, $p = .06$). It also shows a marginal by items site*phase interaction (F_1 (1, 6) = 2.44, MSe = 805.37, $p = .17$; F_2 (1, 56) = 3.61, MSe = 1875.80, $p = .06$; also, phase: F_1 (1, 6) = 8.52, MSe = 1103.12, $p < .05$; F_2 (1, 56) = 15.10, MSe = 2531.64, $p < .001$; rest of *p*s, $p > .25$).

Given the above interactions, further analysis concentrated on the **conditions after NCBV TMS.**

2-way ANOVAs: A two-way ANOVA for the conditions **after NCBV TMS (priming type, relatedness)** showed **a clear priming type*relatedness interaction (F_1 (1, 6) = 6.03, MSe = 363.22, $p < .05$; F_2 (1, 112) = 5.90, MSe = 3174.45, $p < .05$**; also, relatedness: F_1 (1, 6) = 4.57, MSe = 363.22, $p = .08$; F_2 (1, 112) = 1.4, $p > .3$).

2-way ANOVAs: On the contrary, two-way ANOVAs for conditions **before NCBV TMS** (priming type*relatedness: F_1 (1, 6) = 3.08, MSe = 239.17, p = .13; F_2 < 1), **before PLCB TMS** (priming type*relatedness: F_1 (1, 6) = 2.49, MSe = 312.72, p = .17; F_2 < 1), and **after PLCB TMS** (priming type*relatedness: F_1 (1, 6) = 3.40, MSe = 57.46, p = .12; F_2 < 1) showed at most marginal effects.

2-way ANOVAs: A comparison of RTs **for the associative priming set in NCBV conditions (phase, relatedness)** demonstrated **a phase*relatedness interaction, which was significant by subjects, and marginal by items (F_1 (1, 6) = 7.64, MSe = 271.86, p < .05; F_2 (1, 56) = 3.40, p = .07**; also, phase: F_1 (1, 6) = 2.74, MSe = 1133.92, p = .15; F_2 (1, 56) = 4.95, MSe = 2046.34, p < .05; relatedness: F_1 (1, 6) = 6.39, MSe = 204.19, p < .05; F_2 (1, 56) = 3.41, MSe = 3600.41, p = .07).

2-way ANOVAs: A comparison of **related items for NCBV conditions (priming type, phase)** showed **a significant by-items and marginal by subjects phase*priming type interaction (F_1 (1, 6) = 4.35, MSe = 507.14, p = .08; F_2 (1, 56) = 8.01, MSe = 1221.84, p < .01**; phase: F_1 (1, 6) = 1.80, MSe = 1516.93, p > .2; F_2 (1, 56) = 5.93, MSe = 1221.84, p < .05).

2-way ANOVAs: A comparison of **unrelated items after TMS (site, priming type) s**howed an unexpected site*priming type interaction, significant by items, and marginal by subjects (F_1 (1, 6) = 5.48, MSe = 68.64, p = .06; F_2 (1, 56) = 4.36, MSe = 1795.01, p < .05; site: F_1 (1, 6) = 1.4, MSe = 683.33, p > .3; F_2 (1, 56) = 2.12, MSe = 1795.01, p = .15).

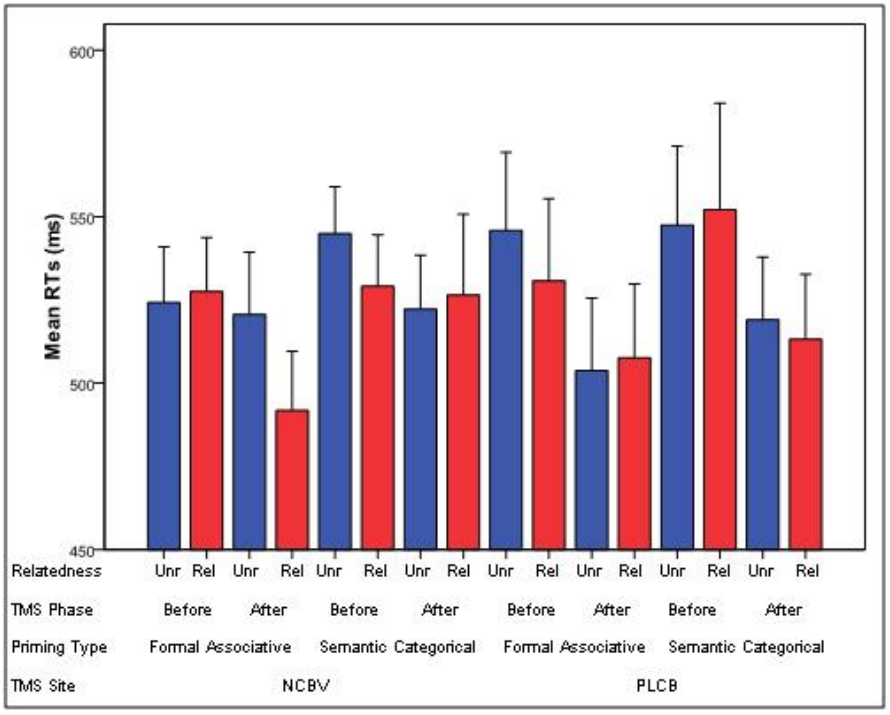
2-way ANOVAs: A comparison of RTs for **related items after TMS (site, priming type)** demonstrated **a significant by-items site*priming type interaction (F_1 (1, 6) = 3.31, MSe = 518.29, p = .12; F_2 (1, 56) = 4.58, MSe = 1457.43, p < .05**; no difference between the two sites: both *F*s, F < 1).

2-way ANOVAs: A comparison of RTs for **associative priming after TMS (site, relatedness)** showed **a site*relatedness interaction (F_1 (1, 6) = 21.19, MSe = 101.690, $p < .005$; F_2 (1, 56) = 5.37, MSe = 1865.18, $p < .05$**; relatedness: F_1 (1, 6) = 3.90, MSe = 318.46, $p = .1$; F_2 (1, 56) = 3.08, MSe = 2879.84, $p = .09$; site: $F < 1$, or $p > .2$).

2-way ANOVAs: A comparison of RTs for the **semantic-categorical** set **after TMS (site, relatedness)** shows no effect of site (site: both *F*s, $F < 1$, or $p > .2$) or of relatedness (both *F*s, $F < 1$), but only a site*relatedness interaction, which was only marginal by items (F_1 (1, 6) = 1.2, MSe = 184.86, $p > .3$; F_2 (1, 56) = 3.54, MSe = 1387.26, $p = .07$).

1-way ANOVAs: formal-associative priming (relatedness): The above interactions were owed to strong **associative priming** occurring **after NCBV TMS (relatedness: F_1 (1, 6) = 10.34, MSe = 322.61, $p < .05$; F_2 (1, 56) = 5.43, MSe = 3476.05, $p < .05$).** On the contrary, for the other conditions, the number of participants did not suffice to show priming (**Associative priming before NCBV TMS:** both *F*s, $F < 1$; **Associative priming before PLCB TMS:** both *p*s, $p > .2$; **Associative priming after PLCB TMS:** both *F*s, $F < 1$).

1-way ANOVAs: semantic-categorical priming (relatedness): Similarly, the number of participants did not suffice to show semantic-categorical priming for any condition (**after NCBV TMS:** both *F*s, $F < 1$, or $p > .25$; **before NCBV TMS:** a nonsignificant trend by subjects for priming: F_1 (1, 6) = 3.06, MSe = 327.86, $p = .13$; F_2 (1, 56) = 1.53, MSe = 1947.81, $p > .2$; **before PLCB TMS:** both *F*s < 1; **after PLCB TMS:** both *F*s, $p > .3$, or $F < 1$).



Plot A.1: Mean latencies for unrelated and related pairs for TMS sessions per condition. Error bars represent + 1 SEM (experiment 3).

TMS Site	Priming Type	TMS Phase	Relatedness	Mean RTs (ms)	SD (ms)
NCBV	**Formal-Associative**	**Before TMS**	**Unrelated**	524.23	47.11
			Related	527.57	45.78
		After TMS	**Unrelated**	520.62	52.80
			Related	491.74	50.64
	Semantic-Categorical	**Before TMS**	**Unrelated**	544.94	39.98
			Related	529.10	43.80
		After TMS	**Unrelated**	522.26	45.72
			Related	526.47	68.60
PLCB	**Formal-Associative**	**Before TMS**	**Unrelated**	545.92	66.23
			Related	530.77	69.60
		After TMS	**Unrelated**	503.72	61.56
			Related	507.66	62.52
	Semantic-Categorical	**Before TMS**	**Unrelated**	547.50	67.19
			Related	552.08	90.50
		After TMS	**Unrelated**	519.06	53.24
			Related	513.12	55.40

Table A.6: Mean latencies for unrelated and related pairs for TMS sessions per condition (experiment 3).

Analysis of accuracy rates for the first TMS study

a) Dependent measure: Priming size (difference between arcsine-transformed mean accuracy rates for unrelated pairs and related ones per condition; see section 5.4.6.1.2).

Mean accuracy rates are shown below in plot A.2 and table A.7 for convenience.

In a **three-way ANOVA** (site, priming type, phase), no effects reached significance without including a between-subjects variable of session number (all *p*s, $p > .2$). This is conceivably because of basic learning effects on the presented words occurring in the second session.

4-way ANOVA (site, priming type, phase, session number): An analysis including 'session number' as a between-subjects variable showed that accuracy differences for the formal-associative set were larger than those for the semantic-categorical set (priming type: F (1, 4) = 5.35, MSe = 0.00, $p = .08$). Interestingly, there was a significant priming type*phase*session number interaction (F (1, 4) = 8.31, MSe = 0.00, $p < .05$). A number of other interactions did not approach significance (priming type*site*phase*session number: F (1, 4) = 2.76, MSe = 0.00, $p = .17$; priming type*site*session number: F (1, 4) = 3.20, MSe = 0.00, $p = .15$; site*session number: F (1, 4) = 3.80, MSe = 0.00, $p = .12$; rest of *p*s, $p > .2$). Subsequent analysis demonstrated that the priming type*phase*session number interaction was due to a rather selective decrease of accuracy differences for formal-associative priming in the first session after TMS, independently of site, with only marginally stronger effects demonstrated after NCBV TMS.

Thus, in order to assess whether that change represents a TMS effect or not, a comparison was performed between the TMS group (n = 8) and the pilot group that

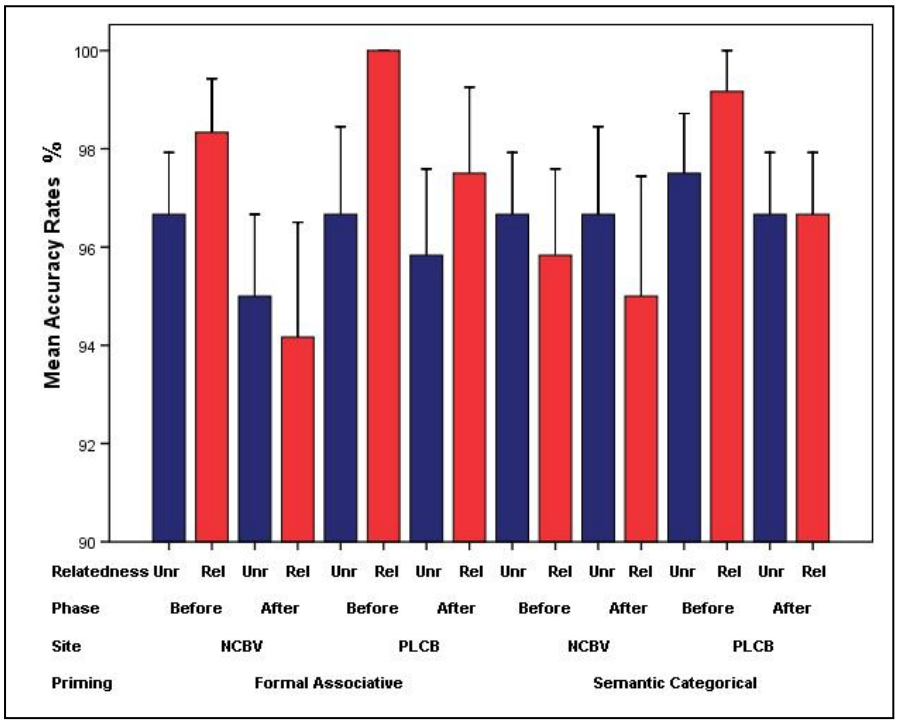
completed two sessions with no TMS (n = 8; section 5.3), collapsing on different site conditions (four-way ANOVA: group, priming type, phase, session number). The comparison demonstrated no priming type*group*phase interaction ($F < 1$). It only showed a significant difference in priming types across conditions, with larger accuracy differences for formal-associative priming than semantic-categorical priming (priming type: F (1, 12) = 5.00, MSe = 0.05, $p < .05$). The analysis also showed a marginal phase*group interaction across priming types (F (1, 12) = 4.06, MSe = 0.00, $p = .07$), and a marginal session number*priming type*group interaction across phases (F (1, 12) = 3.78, MSe = 0.00, $p = .08$). The rest of the interactions did not approach significance (session number*group: F (1, 12) = 2.21, MSe = 0.00, $p = .16$; session number*priming type*phase: F (1, 12) = 2.91, MSe = 0.00, $p = .11$; rest of Fs, $F < 1$). Another such comparison was attempted, this time between the subjects that underwent NCBV TMS in their first session (n = 4) and the no TMS pilot group (n = 8); however, no differences were shown whatsoever (all ps, $p > .2$).

b) Dependent measure: arcsine-transformed accuracy rates

A four-way ANOVA (priming type, TMS site, TMS phase, relatedness) was also performed on raw, arc-sine transformed accuracy rates. The analysis only showed a significant by-subjects and marginal by items decrease in accuracy rates in the second phase (phase: F_1 (1, 6) = 10.64, MSe = 0.00, $p < .05$; F_2 (1, 112) = 3.05, MSe = 0.05, $p = .08$); it also showed an unexpectedly significant by items difference in accuracy rates between the two sites overall, with higher accuracy rates for PLCB conditions (site: F_1 (1, 6) = 1.80, MSe = 0.00, $p > .25$; F_2 (1, 112) = 4.09, MSe = 0.05, $p < .05$; no phase*relatedness interaction: F_1 (1, 6) = 2.20, MSe = 0.00, $p > .2$; F_2 (1, 112) = 2.37, MSe = 0.05, $p = .13$; no site*phase*priming type interaction: F_1 (1, 6) = 3.84, MSe = 0.00, $p = 0.1$; $F_2 < 1$; rest of ps, $p > .2$).

Priming Type	TMS Site	TMS Phase	Relatedness	Mean accuracy rates (%)	SD (%)
Formal-Associative	NCBV	Before TMS	Unrelated	97	4
			Related	98	3
		After TMS	Unrelated	95	5
			Related	94	7
	PLCB	Before TMS	Unrelated	97	5
			Related	100	0
		After TMS	Unrelated	96	5
			Related	98	5
Semantic-Categorical	NCBV	Before TMS	Unrelated	97	4
			Related	96	5
		After TMS	Unrelated	97	5
			Related	95	7
	PLCB	Before TMS	Unrelated	98	3
			Related	99	2
		After TMS	Unrelated	97	4
			Related	97	4

Table A.7 : Mean accuracy rates for unrelated and related pairs for TMS sessions per condition (experiment 3).



Plot A.2 : Mean accuracy rates for unrelated and related pairs for TMS sessions per condition. Error bars represent + 1 SEM (experiment 3).

Appendix B

Material for Experiment 4

(TMS Study 2)

Formal-Associatively Related Pairs			
Prime	**Translation**	**Target**	**Translation**
φίδι	snake	τρύπα	hole
κοντσέρτο	concert	πολυβόλα	machine-guns
άρτος	bread	θεάματα	spectacles
παρτίδα	game'	σκάκι	chess
κρασί	wine	εικοσιένα	twenty-one (n)
ζευγάρι	pair	παπούτσια	shoes
μήλο	apple	κανέλα	cinnamon
δολοφόνος	killer	πριόνι	saw
έρωτας	love	χιόνια	snow
γροθιά	punch	στομάχι	stomach
δανεισμός	loan	κρυφά	in hiding'
άγιος	saint	φοβέρα	scaring'
γούστο	taste	καπέλο	cap
τσάι	tea	λεμόνι	lemon
χήρα	widow	κρεβάτι	bed
αγκάθι	spike	πλευρό	side
ακτίνες	rays	γάμμα	gamma
τρελός	crazy (n)	δέσιμο	tying
τσάντα	bag	κοπάνα	'school-avoiding'
τρίποντο	'three point-shot'	εκπνοή	end
λουκέτο	locket	μαγαζιά	stores
γάτα	cat	πέταλα	petals
έπος	epic	σαράντα	forty
κοριτσάκι	girlie	σπίρτα	matches
πυρ	fire	μανία	mania
σκατό	shit	παξιμάδι	cracker
λίστα	list	ψώνια	shopping
ποτάμι	river	δάκρυα	tears
κολοκύθια	pumpkins	ρίγανη	oregano
κουτί	box	γλυκά	sweets
ζήλεια	jealousy	ψώρα	itch
καβούρια	crabs	τσέπες	pockets
βλέμμα	sight	κάλπη	ballot box
γυαλί	glass	φαρμάκι	poison
ρώγα	nipple	σταφύλι	grape
σκόνη	dust	θρύψαλα	shivers
δωμάτιο	room	θέα	view
μπαμπάδες	'babades'(sweet)	ρούμι	rum
αλλεργία	allergy	γάλα	milk
κουνούπι	mosquito	τίγρης	tiger
βίρα	'lift (the anchors)'(exclamation)	άγκυρες	anchors
τόλμη	boldness	γοητεία	charm

στουπί	tow	μεθύσι	inebriation
στρίγκλα	shrew	αρνάκι	lamb
γεύση	taste (n)	φράουλα	strawberry
γουρούνι	pig	σακί	sack
ένοχος	guilty (n)	αιτία	cause
λατέρνα	lantern	φτώχεια	poverty
φτωχός	poor (n)	μοίρα	fate
φιδάκι	snake	διαμαντής	‘Diamantis’ (name)
σκατά	shit	μούτρα	face
ουρανός	sky	σφοντύλι	spindle
άσος	ace	μανίκι	sleeve
βαπόρι	small boat	περσία	Persia
κοτόπουλο	chicken	σχάρα	grill
νταντά	nanny	ζόρι	strain
χταπόδι	octopus	κάρβουνα	coal
ράμματα	stitches	γούνα	fur
μάτσο	bunch	χάλια	mess
δέοντα	‘appropriate things'	μαμά	mama
ναύτης	sailor	στεριά	land
τόπι	ball	ξύλο	wood (beating)
κάρο	cart	μαλακίες	‘wanking' (‘bullocks’)
μάγισσες	witches	πυρά	fire
φασόλια	beans	γίγαντες	giants
τρένο	train	φάντασμα	phantom
καιρός	weather	χάσιμο	loss
θάνατος	death	αφεντικά	bosses
κρίμα	shame	νιάτα	youth
βουτιά	dive	κενό	void
αυγά	eggs	πασχάλια	lilac
μαϊμού	monkey	δελφίνι	dolphin
αχλάδα	pear	ουρά	tale
φέτα	slice	ψωμί	bread
σήμα	sign	λιοντάρι	lion
ουρά	tail	σκέλια	legs
ελιά	olive	μάγουλο	cheek
αέρας	wind	πανιά	sales
γκάζι	gas	πάτωμα	floor
φύκια	sea-weed	κορδέλες	ribbons
αλήτες	punks	πουλιά	birds
γαρύφαλλο	carnation	αυτί	ear
διάολος	devil	ποδάρι	leg
βάλσαμο	balsam	πληγές	wounds
βελόνα	needle	άχυρα	hay
κραυγές	cries	σκοτάδι	darkness

ψύλλοι	fleas	αυτιά	ears
ποντίκι	mouse	φάκα	trap
λεφτά	money	πέταμα	'throwing away'
κάστανα	chestnut	φωτιά	fire
κόντρα	'counter' (shaving method)	ξύρισμα	shaving
θηρίο	beast	κλουβί	cage
κέρατο	horn	σύννεφο	cloud
πρίγκηπας	prince	βάτραχος	frog
ψάρι	fish	χείλη	lips
πρόβατο	sheep	σφαγή	slaughter
απίδια	pears	σάκος	sack
γορίλες	gorillas	ομίχλη	mist
στεφάνι	wreath	μνημείο	monument
γόβα	court shoe	στιλέτο	dagger

Table B.1. Formal-associatively related pairs (experiment 4).

μπουκάλι	bottle	λεκάνη	basin
ψάρι	fish	βοδινό	beef
φανάρι	beacon	γρασίδι	grass
συμβόλαιο	contract	χρέος	debt
νήμα	thread	καλώδιο	cable
βάρκα	boat	σχεδία	raft
κοπριά	dung	λάσπη	mud
λίμνη	lake	πέλαγος	sea
τρίψιμο	rubbing	ξέπλυμα	washing
γκρεμός	precipice	ράχη	back
δειλινό	evening	μεσάνυχτα	midnight
αυλή	yard	χωράφι	field
πηλός	clay	τσιμέντο	cement
βότσαλο	pebble	λίθος	stone
σπυράκια	acme	ελιές	mole
κόκκοι	grains	σταγόνες	drops
ψωμί	bread	πίτα	pie
δέμα	parcel	κάρτα	card
σούπα	soup	γλυκό	dessert
προδότης	trator	καρφί	nail
καημός	distress	παράπονο	complaint
μάνα	mother	θείος	uncle
μυς	muscle	ιστός	tissue
νάνος	dwarf	μάγισσα	witxh
έρημος	desert	ζούγκλα	jungle
βεβαίωση	certificate	δίπλωμα	diploma
τίτλοι	titles	μετρητά	cash
βοσκός	shepherd	αγρότης	farmer
αγωνία	‘agony’ (uno, card game)	ξερή	‘kseri’ (card game)
πετάλι	pedal	φρένο	break
ντάμα	queen	άσος	ace
στίχος	verse	αράδα	line
σταυρός	cross	χάντρα	bead
ποίημα	poem	νουβέλα	novel
σεισμός	earthquake	καταιγίδα	storm
γόνατο	knee	καλάμι	shin
αετός	eagle	γεράκι	hawk
μύγα	fly	κουνούπι	mosquito
έγκλημα	crime	παράβαση	violation
κύλινδρος	cylinder	σφαίρα	sphere
πάλη	wrestling	στίβος	athletics
έλικες	screw	φτερά	feathers
ύψη	heights	φίδια	snakes
μάρτυρας	witness	άγγελος	angel

πέταλο	petal	θύρα	gate
βάρη	weights	κάμψεις	crunches
κάταγμα	fracture	ρήξη	laceration
εξοχικό	villa	καλύβα	cottage
βαρύτητα	gravity	άνωση	buoyancy
σύρραξη	conflict	αγώνας	fight
μήλα	apples	κουτσό	lame' (child game)
μπαταρία	battery	λάστιχο	tire
κράνος	helmet	φέσι	fez
φωλιά	nest	μαντρί	kraal
πλώρη	prow	κατάρτι	mast
καρκίνος	cancer	γρίπη	flu
πίνακας	board	γλυπτό	sculpture
δεξαμενή	tank	κιβώτιο	case
ακτίνα	ray	τόξο	bow
ράβδος	bar	ακόντιο	spear
κόψιμο	cut	πόνος	pain
μούσα	muse	νύμφη	nymph
μαξιλάρι	pillow	ντιβάνι	mattress
γένος	gender	φυλή	tribe
σεντόνι	sheet	κουβέρτα	blanket
τροχός	wheel	ένεση	shot
πίπα	pipe	τσιγάρο	cigarette
λεξικό	dictionary	βίβλος	bible
φρουρός	guard	ζητάς	cop'
νησί	island	ήπειρος	continent
τσούρμο	pack	παρέα	company
λόφος	hill	ηφαίστειο	volcano
κατάθεση	deposit	δόσεις	flop
εμβόλιο	vaccine	χάπι	pill
οικοδομή	structure	λιμάνι	port
κόρνα	horn	στέρεο	stereo
δαγκάνες	mandibles	νύχια	nails
κοπάδι	flock	σμήνος	raft
φόνος	murder	κλοπή	theft
σέλα	saddle	χαλινάρια	gains
χούντα	dictatorship	μοναρχία	monarchy

Semantic-Categorically Related Pairs			
Prime	**Translation**	**Target**	**Translation**
ρήμα	verb	επίθετο	adjective
κερί	candle	λάμπα	lamp
ηρωίνη	heroine	έκσταση	ecstasy
κλαδί	branch	φύλλο	leaf
τυρί	cheese	κιμάς	mince
τάγμα	battalion	μεραρχία	division
τόνος	tuna	κιλό	kilogram
τζάμπα	in vein'	στράφι	'waste'
εστία	goal	καλάθι	basket
σακάκι	jacket	μπέρτα	cape
σίδερο	iron	χρυσός	gold
μπαράκι	bar	ταβέρνα	tavern
πιατέλα	plateau	δίσκος	discus
μπαλέτο	ballet	τανγκό	tango
συμμορία	gang	κύκλωμα	circuit
φράγκα	franks	λίρες	pounds
οξυγόνο	oxygen	κάλιο	calcium
πλέγμα	fence	μπάρα	bar
πρίγκιπας	prince	λόρδος	lord

Table B.2 : Semantic-categorically related pairs (experiment 4).

Measures	**Priming Set**	**Mean**	**SD**
Prime word HNC frequency (words/million)	**Associative**	7.08	9.89
	Coordinate	7.43	10.30
	Total	7.25	10.08
Prime word length (number of letters)	**Associative**	6.02	1.52
	Coordinate	6.00	1.43
	Total	6.01	1.47
Target word HNC frequency (words/million)	**Associative**	7.48	10.27
	Coordinate	7.11	9.41
	Total	7.30	9.82
Target word length (number of letters)	**Associative**	6.03	1.32
	Coordinate	6.01	1.28
	Total	6.02	1.30

Table B.3: Measures for formal-associatively and semantic-categorically related items (experiment 4).

Measures	List	Mean	SD
Prime word HNC frequency (words/million)	**1**	7.16	9.74
	2	7.34	10.44
	Total	7.25	10.08
Prime word length (number of letters)	**1**	5.97	1.47
	2	6.05	1.48
	Total	6.01	1.47
Target word length (number of letters)	**1**	5.95	1.42
	2	6.09	1.16
	Total	6.02	1.30
Target word HNC frequency (words/million)	**1**	7.26	9.77
	2	7.34	9.93
	Total	7.30	9.82

Table B.4: Word measures (experiment 4).

Measures	List	Mean	SD
Prime word HNC frequency (words/million)	**1**	7.51	7.91
	2	7.52	6.62
	Total	7.52	7.22
Prime word length (number of letters)	**1**	5.56	1.69
	2	5.88	1.33
	Total	5.72	1.51
Target word length (number of letters)	**1**	5.84	1.03
	2	6.04	1.46
	Total	5.94	1.25

Table B.5: Nonword measures (experiment 4).

Appendix C

Test Items for experiment 5

(TMS study 3)

Semantic-Categorically related		Formal-Associatively related		Semantic-Associatively related	
Prime	**Target**	**Prime**	**Target**	**Prime**	**Target**
trick	deceiving	trail	blazing	president	governing
clash	disagreeing	bone	chilling	snake	devouring
pardon	forgiving	sugar	coating	burglar	stealing
collection	gathering	dynamite	fishing	merchant	selling
damage	hurting	fortune	telling	thief	robbing
sack	firing	finger	licking	carpenter	hammering
curse	swearing	peace	offering	rapist	torturing
prowl	lurking	nerve	racking	platter	serving
compliance	obeying	figure	skating	furnace	heating
comment	remarking	proof	reading	dagger	stabbing
vision	seeing	echo	locating	garage	parking
selection	picking	globe	trotting	studio	drawing
waste	spending	type	setting	racetrack	betting
lecture	teaching	stone	walling	park	strolling
permission	allowing	fire	watching	juice	drinking
crush	smashing	mud	wrestling	crook	arresting
arson	burning	hang	gliding	nominee	presenting
suck	absorbing	sun	bathing	teeth	brushing
persuasion	convincing	shop	lifting	dictator	overthrowing
stress	emphasizing	steeple	chasing	friend	accepting
blast	exploding	salad	dressing	spectator	cheering
repair	fixing	marriage	counselling	stripper	entertaining
risk	gambling	mountain	biking	maid	cleaning
audition	hearing	soul	destroying	nun	praying
aid	helping	nose	diving	hose	spraying
assessment	marking	eaves	dropping	towel	drying
failure	missing	navel	gazing	laboratory	examining
guess	predicting	fun	loving	theatre	acting
complaint	protesting	shotgun	wedding	cemetery	mourning
transmission	sending	hair	splitting	library	studying
formation	shaping	wire	tapping	taxpayer	auditing
lapse	slipping	heart	warming	cashier	paying
loot	plundering	ghost	writing	bird	startling
journey	travelling	combat	training	cattle	transporting
expansion	growing	street	fighting	athlete	competing
extension	prolonging	stage	managing	knife	slicing
entrance	accessing	tongue	lashing	instructor	evaluating
breath	inhaling	baby	sitting	rag	dusting
ignorance	neglecting	blood	sucking	chainsaw	cutting
expulsion	ousting	role	modelling	gardener	scratching

reply	answering	spell	binding	parent	adopting
relation	associating	eye	catching	terrorist	executing
hypothesis	assuming	ground	breaking	lawyer	questioning
rush	surging	panic	buying	scientist	considering
offence	attacking	winter	clothing	host	inviting
refusal	declining	belly	dancing	lover	kissing
finish	ending	plea	bargaining	reporter	interviewing
detection	finding	treasure	hunting	hammer	repairing
leap	jumping	contract	killing	needle	sewing
retention	keeping	emergency	landing	canoe	floating
shortage	lacking	air	conditioning	keyboard	typing
rent	letting	hell	raising	kitchen	cooking
command	ordering	joy	riding	synagogue	preaching
defence	protecting	steam	rolling	restaurant	eating
stay	resting	heat	seeking	bedroom	sleeping
discovery	revealing	file	sharing	inmate	releasing
vibration	shaking	chain	smoking	egg	frying
prevention	stopping	flame	throwing	pupil	dismissing
clap	applauding	safety	warning	puzzle	solving
gulp	swallowing	brain	washing	clothes	wearing
emergence	appearing	piggy	backing	prosecutor	accusing
request	asking	tree	hugging	actor	performing
arrest	capturing	sand	blasting	boss	hiring
gossip	chatting	shadow	boxing	nurse	identifying
desire	craving	stock	broking	specialist	requesting
fear	dreading	gate	crashing	priest	worshipping
departure	leaving	sleep	walking	oxygen	reviving
emancipation	liberating	share	cropping	brush	painting
walk	marching	zebra	crossing	ruler	measuring
cause	provoking	bottle	feeding	museum	browsing
attraction	pulling	gut	feeling	ocean	swimming
blow	punching	test	driving	airport	flying
arrival	reaching	student	funding	beach	tanning
ablation	removing	gold	plating	celebrity	recognising
salvation	rescuing	trouble	shooting	movie	renting
spin	revolving	window	shopping	newborn	carrying
gaze	staring	code	switching	candidate	choosing
rap	striking	cream	filling	prey	stalking
proposal	suggesting	rote	learning	fan	admiring
intrusion	violating	mouth	watering	napkin	wiping

Table C.1 : Semantic-categorically, formal-associatively, and semantic-associatively related pairs (experiment 5).

Measures	Mean	SD
Prime word BNC frequency (words/million)	36.05	35.18
Prime word length (letters)	6.78	2.30
Target word BNC frequency (words/million)	19.64	25.30
Target word length (letters)	8.18	1.30
Semantic Similarity (Wordnet::Similarity Ratings)	0.74	0.25
BNC phrase frequency (phrases/million)	0.00	0.00

Table C.2: Measures for semantic-categorically related items (experiment 5).

Measures	Mean	SD
Prime word BNC Frequency (words/million)	25.52	34.39
Prime word length (letters)	6.48	1.87
Target word BNC frequency (words/million)	15.65	17.16
Target word length (letters)	8.30	1.62
Semantic Similarity (Wordnet::Similarity ratings)	0.12	0.06
BNC phrase frequency (phrases/million)	0.05	0.20

Table C.3: Measures for semantic-associatively related items (experiment 5).

Measure	Mean	SD
Prime word BNC frequency (words/million)	65.18	58.03
Prime word length (letters)	5.04	1.34
Target word BNC frequency (words/million)	30.28	37.70
Target word length (letters)	7.48	1.04
Semantic Similarity (Wordnet::Similarity ratings)	0.18	0.07
BNC phrase frequency (phrases/million)	0.35	0.46

Table C.4: Measures for formal-associatively related items (experiment 5).

Measures	List	Mean	SD
Prime Word BNC Frequency (words/million)	**A**	30.67	48.26
	B	30.21	22.34
	Total	30.44	37.52
Prime Word Length (number of letters)	**A**	5.84	1.78
	B	5.86	2.19
	Total	5.85	1.99
Target Word Length (number of letters)	**A**	8.26	1.49
	B	8.22	1.49
	Total	8.24	1.49

Table C.5: Measures for nonword trials across the two lists (experiment 5).

Appendix D

Material for Experiments 6 and 7

(TMS Study 3)

Semantic-Categorically Related		Semantic-Associatively Related		
prime	**target**	**θ-role**	**prime**	**target**
reply	answering	Agent	fan	admiring
emergence	appearing	Agent	friend	accepting
relation	associating	Agent	prosecutor	accusing
offence	attacking	Agent	parent	adopting
arrest	capturing	Agent	committee	approving
spread	distributing	Agent	spectator	cheering
finish	ending	Agent	maid	cleaning
sack	firing	Agent	athlete	competing
repair	fixing	Agent	musician	composing
pardon	forgiving	Agent	scientist	considering
expansion	growing	Agent	snake	devouring
damage	hurting	Agent	stripper	entertaining
leap	jumping	Agent	instructor	evaluating
shortage	lacking	Agent	terrorist	executing
emancipation	liberating	Agent	president	governing
failure	missing	Agent	carpenter	hammering
command	ordering	Agent	boss	hiring
expulsion	ousting	Agent	reporter	interviewing
selection	picking	Agent	host	inviting
defence	protecting	Agent	lover	kissing
discovery	revealing	Agent	policeman	lecturing
formation	shaping	Agent	surgeon	operating
crush	smashing	Agent	actor	performing
rap	striking	Agent	nun	praying
receipt	taking	Agent	babysitter	punishing
permission	allowing	Agent	lawyer	questioning
clap	applauding	Agent	philosopher	reading
pursuit	chasing	Agent	singer	recording
persuasion	convincing	Agent	specialist	requesting
fear	dreading	Agent	thief	robbing
anticipation	expecting	Agent	merchant	selling
blast	exploding	Agent	sniper	shooting
perception	feeling	Agent	salesman	showing
detection	finding	Agent	pitcher	throwing
collection	gathering	Agent	rapist	torturing
aid	helping	Agent	priest	worshipping
retention	keeping	Instrument	stove	cooking
departure	leaving	Instrument	chainsaw	cutting
walk	marching	Instrument	towel	drying
assessment	marking	Instrument	rag	dusting
loot	plundering	Instrument	rod	fishing
guess	predicting	Instrument	canoe	floating

attraction	pulling	Instrument	furnace	heating
blow	punching	Instrument	ruler	measuring
spin	revolving	Instrument	brush	painting
vibration	shaking	Instrument	hammer	repairing
rush	surging	Instrument	oxygen	reviving
gulp	swallowing	Instrument	platter	serving
curse	swearing	Instrument	needle	sewing
experience	tasting	Instrument	knife	slicing
refusal	declining	Instrument	hose	spraying
prowl	lurking	Instrument	dagger	stabbing
murder	killing	Instrument	keyboard	typing
ignorance	neglecting	Instrument	detergent	washing
comment	remarking	Instrument	binoculars	watching
cause	provoking	Instrument	napkin	wiping
rent	letting	Location	theatre	acting
gaze	staring	Location	racetrack	betting
arrival	reaching	Location	studio	drawing
journey	travelling	Location	highway	driving
victory	winning	Location	restaurant	eating
hypothesis	assuming	Location	laboratory	examining
lecture	teaching	Location	airport	flying
request	asking	Location	forest	hunting
entrance	accessing	Location	cemetery	mourning
practice	exercising	Location	garage	parking
ablation	removing	Location	synagogue	preaching
erection	lifting	Location	bathroom	showering
proposal	suggesting	Location	bedroom	sleeping
quest	seeking	Location	library	studying
complaint	protesting	Location	ocean	swimming
audition	hearing	Location	beach	tanning
salvation	rescuing	Patient	crook	arresting
clash	disagreeing	Patient	taxpayer	auditing
purchase	buying	Patient	teeth	brushing
arson	burning	Patient	newborn	carrying
gossip	chatting	Patient	cheque	cashing
lapse	slipping	Patient	candidate	choosing
trick	deceiving	Patient	hair	combing
compliance	obeying	Patient	patient	curing
intrusion	violating	Patient	pupil	dismissing
waste	spending	Patient	juice	drinking
prevention	stopping	Patient	baby	frightening
breath	inhaling	Patient	student	grading
risk	gambling	Patient	wimp	kicking
robbery	stealing	Patient	husband	loving

motion	walking	Patient	gift	opening
vision	seeing	Patient	dictator	overthrowing
stress	emphasizing	Patient	cashier	paying
fraction	breaking	Patient	nominee	presenting
extension	prolonging	Patient	celebrity	recognising
proof	verifying	Patient	bottle	recycling
desire	craving	Patient	inmate	releasing
search	browsing	Patient	movie	renting
suck	absorbing	Patient	gardener	scratching
stay	resting	Patient	puzzle	solving
argument	fighting	Patient	prey	stalking
composition	writing	Patient	bird	startling
expectation	waiting	Patient	cattle	transporting
recognition	identifying	Patient	clothes	wearing

Table D.1 : Semantic-categorically and semantic-associatively related pairs (experiments 6 and 7).

Measures	Minimum	Maximum	Mean	SD
Prime word BNC frequency (words/million)	0.61	234.12	39.89	41.69
Prime word length (letters)	3.00	12.00	6.97	2.25
Target word BNC frequency (words/million)	0.60	240.57	26.05	37.16
Target word length (letters)	6.00	12.00	8.12	1.36
Semantic Similarity (Wordnet::Similarity Ratings)	0.50	1.00	0.73	0.25
BNC phrase frequency (phrases/million)	0.00	0.01	0.00	0.00

Table D.2: Semantic-categorically related items (experiments 6 and 7).

Measures	Minimum	Maximum	Mean	SD
Prime word BNC frequency (words/million)	0.67	209.64	27.58	39.70
Prime word length (letters)	3.00	11.00	6.61	1.92
Target word BNC frequency (words/million)	0.66	119.57	19.41	22.81
Target word length (letters)	6.00	12.00	8.15	1.51
Semantic Similarity (Wordnet::Similarity Ratings)	0.05	0.33	0.12	0.06
BNC phrase frequency (phrases/million)	0.00	0.02	0.00	0.00

Table D.3: Semantic-associatively related items (experiments 6 and 7).

Measures	Pair	Mean	SD
Prime word length (number of letters)	Word-Nonword	6.66	1.43
	Word-Word	6.78	1.92
	Total	6.72	1.69
Target word length (number of letters)	Word-Nonword	8.19	1.34
	Word-Word	8.14	1.53
	Total	8.17	1.44
Prime word BNC frequency (words/million)	Word-Nonword	32.70	39.93
	Word-Word	33.62	40.83
	Total	33.16	40.34

Table D.4: Measures for word-word and word-nonword pairs (experiments 6 and 7).

Appendix E

TMS Safety Declaration Form

TMS Studies 1 and 3

TMS SAFETY DECLARATION

To be completed before testing commences

Please answer the following confidential questions by circling YES or NO to each one. It is important that you answer these questions accurately as some of the items may be of importance to your safety.

Do you have a history of epilepsy and/or seizures?

YES/NO

Have any of your first degree relatives had a history of epilepsy or seizures?

YES/NO

Do you have a cardiac pacemaker fitted?

YES/NO

Have you had any operations on your head?

YES/NO

Do you have any joint replacements or metal implants in any of your body or head?

YES/NO

Are you taking any medication that could influence the nervous system, such as anti-depressants, anti-epileptics, or sleeping pills?

YES/NO

Have you participated in any other TMS experiments earlier today?

YES/NO

TO BE ANSWERED BY WOMEN OF A CHILD BEARING AGE

Could you be pregnant?

YES/NO

I confirm that I have read the above questions and that my answers are correct to the best of my knowledge and belief

Signature

Name (BLOCK CAPITALS)

Date

Appendix F

TMS Safety Declaration Form

TMS Study 2

ΕΜΠΙΣΤΕΥΤΙΚΕΣ ΕΡΩΤΗΣΕΙΣ

για πείραμα Διακρανικού Μαγνητικού Ερεθισμού (ΔΜΕ)

Προς συμπλήρωση απο τους συμμετέχοντες

Παρακαλώ απαντήστε τις παρακάτω εμπιστευτικές ερωτήσεις κυκλώνοντας ΝΑΙ ή ΟΧΙ για την κάθε μια. Διαβάστε τις ερωτήσεις με προσοχή, καθώς αφορούν άμεσα στην ασφάλειά σας.

1. Εχετε **εσείς ή οποιοσδήποτε πρώτου βαθμού συγγενής (εφόσον γνωρίζετε) ιστορικό επιληπτικών κρίσεων;**

ΝΑΙ/ΟΧΙ

2. Φέρετε οποιουδήποτε τύπου **καρδιακό βηματοδότη, προσθετικά μέλη ή μεταλλικά εμφυτεύματα** σε οποιοδήποτε μέρος του σώματός σας;

ΝΑΙ/ΟΧΙ

3. Έχετε υποστεί οποιουδήποτε τύπου **εγχείρηση στον εγκέφαλό** σας;

ΝΑΙ/ ΟΧΙ

4. Λαμβάνετε οποιαδήποτε **φαρμακευτική αγωγή** που θα μπορούσε να επηρεάσει το νευρικό σας σύστημα, όπως αντικαταθλιπτικά, αντιεπιληπτικά, ηρεμιστικά;

ΝΑΙ/ ΟΧΙ

ΠΡΟΣ ΑΠΑΝΤΗΣΗ ΑΠΟ ΓΥΝΑΙΚΕΣ

Ενδέχεται να είστε έγκυος;

ΝΑΙ / ΟΧΙ

Appendix G

Published and Submitted Papers

This appendix contains articles that were published or submitted prior to the completion of the thesis; it consists of one submitted journal article (Argyropoulos, 2010b), two articles published in the proceedings of the 7th and 8th EvoLang conferences (Argyropoulos, 2008a, 2010a), one article published in the first BrainTalk conference (Argyropoulos, 2009), and one journal commentary (Hurford, Flaherty, & Argyropoulos, 2007). Details of these articles are given below:

Argyropoulos, G. P. (2008a). 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. (2010a). Is grammaticalization glossogenetic? In A. D. M. Smith, M. Schouwstra, B. de Boer, & K. Smith (Eds.), *The Evolution of Language: Proceedings of the 8th International Conference on the Evolution of Language* (pp. 3-10). Singapore: World Scientific.

Argyropoulos, G. P. (2010b). *Theta-burst stimulation of the right neocerebellar vermis selectively enhances lexical associative priming.* Manuscript submitted for publication.

Hurford, J., Flaherty, M., & Argyropoulos, G. (2007). Past and future, human and non-human, semantic/procedural and episodic. *Behavioral and Brain Sciences, 30*(3), 324-325. Commentary on T. Suddendorf & M. Corballis *The evolution of foresight: What is mental time travel and is it unique to humans?*

Argyropoulos, G. P. (2008a). 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.

THE SUBCORTICAL FOUNDATIONS OF GRAMMATICALIZATION

GIORGOS P. ARGYROPOULOS

Language Evolution and Computation Research Unit, University of Edinburgh
40 George Square, Edinburgh, EH8 9LL, Scotland, UK
giorgos@ling.ed.ac.uk

The present paper raises the so far unaddressed question of the neurolinguistic processes underlying grammaticalization operations. Two adaptive mechanisms are presented, based on current research on the subcortical contributions to aspects of higher cognition: The cerebellar-induced Kalman gain reduction in linguistic processing, and the basal ganglionic re-regulation of cortical unification operations.

1. Introduction

The neuroanatomy of either the domain-general cognitive phenomena underlying grammaticalization, e.g., "ritualization" (Haiman, 1994), "automatization" (Givón, 1979; Bybee, 1998), or the particular psycholinguistic processes, e.g., Pickering and Garrod's (2004) dialogical "routinization", has hardly attracted any attention in the literature. Haiman's (1994, p.25) comment, that "the physiology of ritualization in human beings is unknown", is rather suggestive. The *desideratum*, then, is to move from the *sine qua non* of the neural grounding of such putative domain-general cognitive phenomena to a neurolinguistics of grammaticalization, by introducing I-language adaptation processes (both representational fine-tuning and executional optimization) in accordance with changing E-language properties.

2. The Explanandum of Grammaticalization

Grammaticalization, "an evolution whereby linguistic units lose in semantic complexity, pragmatic significance, syntactic freedom, and phonetic substance" (Heine & Reh, 1984, p.15), is a manifestation of the "Reducing Effect" of repetition in linguistic behaviour (Bybee & Thompson, 2000): "Univerbation" (Lehmann, 1995), i.e., the gain in syntagmatic bondedness (e.g., *hac hora* (Latin) > *ahora* (Spanish)), "phonetic attrition" (Givón, 1979), i.e., the minimization of articulatory gestures (e.g., going to > gonna),

desemanticization, i.e., the loss of (lexical) meaning of a particular item (e.g., future marker *"will"* loses the meaning of desire), are the fundamental aspects of this process. Because of its desemanticization, the particular item occurs in a greater contextual variety, inviting additional inferences, inducing its "context-induced reinterpretation" (Heine, Claudi, & Hünnemeyer, 1991). As a result, such item behaviorally deviates from its particular category, i.e., it is decategorialized, (e.g., the V > P cline in English).

3. From Grammaticalization to the Cerebellum and the Basal ganglia

Automatization, however, the cognitive basis of grammaticalization, is known to rely on the basal ganglia (BG) and the cerebellum (CB) (e.g., Thach, Mink, Goodkin, & Keating, 2000): Signals from the cerebral cortex are optimized on the basis of their reward value (reinforcement learning) and their accuracy (supervised learning) through the BG and CB loop circuits, respectively (Doya, 1999). The Cerebellar CorticoNuclear MicroComplex (CNMC) (Ito, 1984), i.e., the CB adaptive unit that learns based on error signals, becomes an internal model, an "emulator" (Grush, 2004), with signal-transfer characteristics identical to those of the copied cortical system **(figure 1)**. Maximized reliance of the CB Kalman Filter (Paulin, 1989) on the predictions of an accurate internal model, i.e., low gain of the Kalman regulator (KG), drastically economizes on attentional-executional resources. On the other hand, the BG "sculpting process" (Graybiel, 2000) induces the context-sensitive fluent gating in a "winner-takes-all" fashion of competing motor actions, via inhibition-disinhibition processes.

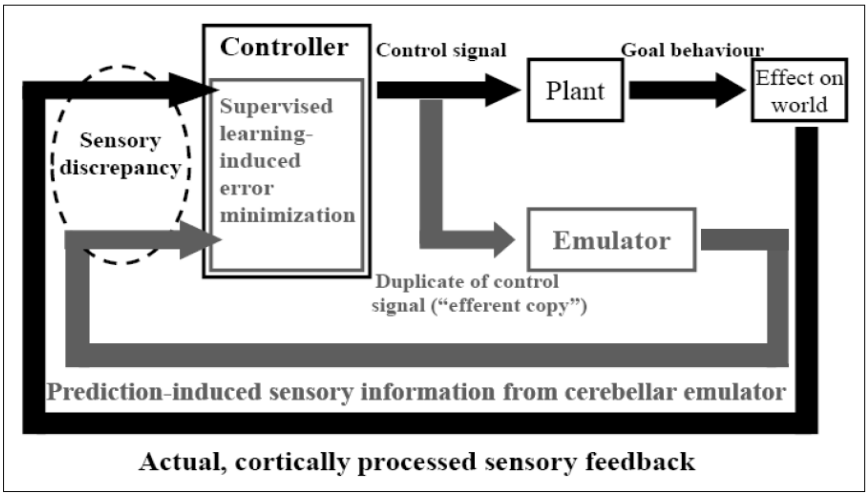


Figure 1. A simple cerebellar feedforward emulator (e.g., Wolpert, Miall, & Kawato, 1998). The predictions of the internal model are constantly updated, based on the error signals of the discrepancy induced by the actual sensory feedback.

4. The Neurolinguistic Grounding of Grammaticalization

I propose that the neurolinguistic basis of the Reducing Effect in grammaticalization is the CB-induced KG reduction in multilevel linguistic processing, and that the one for the formation-deformation of probabilistic categories is the BG adaptive regulation of unification operations.

4.1. *Cerebellar-induced Kalman Gain Reduction in Linguistic Processing*

The CB as a neural analog of a dynamical state estimator (Paulin, 1989) provides a highly plausible basis for Pickering and Garrod's (2007) Kalman filter-processor **(figure 2)**. Suggestively, CB error-signaling is involved in sentence processing (Stowe, Paans, Wijers, & Zwarts, 2004). Lack of performance optimization (interpretable as KG reduction) for CB patients in linguistic tasks is well established (Fiez, Petersen, Cheney, & Raichle, 1992), while in CB aphasiology the notion of "neurofunctional redundancy" has been invoked for the CB (emulated) linguistic representations: CB aphasia is significantly milder than classical aphasic syndromes, owing to maximal prefrontal cortical compensation. (Fabbro *et al*., 2004).

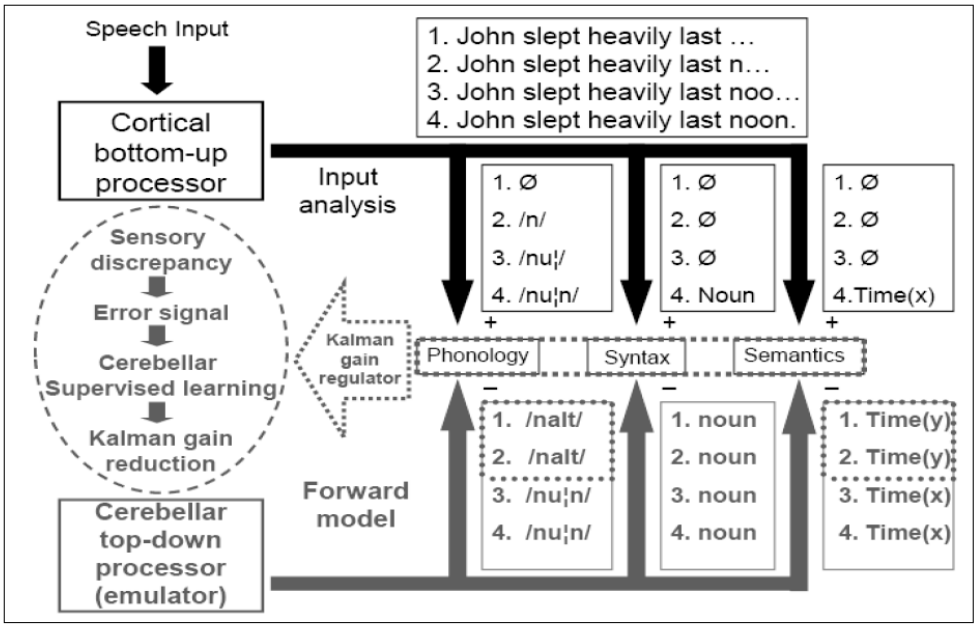


Figure 2. The CB as a domain general Kalman Filter (Paulin 1989) meets Pickering and Garrod (2007): Minimization of the gain of the Kalman regulator via routinization corroborates the reliance on a top-down (expectation-based) processing modality; "shallow processing" (Barton & Sanford, 1993) and "good-enough representations" (Ferreira, Bailey, & Ferraro, 2002) are suggestive cases.

4.1.1. *Chunking and phonetic attrition*

Univerbation is a case of chunking (e.g., Haiman, 1994; Bybee, 1998), i.e., the creation of compound behavioural units the interior of which exhibits minimal attentional and executional costs. Chunking is a well-established CB-induced cognitive function: CB deficits exhibit lack of practice-induced facilitation (e.g., LaForce & Doyon, 2001), and decomposition of motor behaviour (e.g., Thach *et al*., 2000). In the same spirit, phonetic attrition is the linguistic instance of the CB-induced minimization of articulatory stiffness in motor behaviour (e.g., Wolpert, Miall, & Kawato, 1998). Suggestively, Ackermann and Hertrich (2000) emphasize the CB's role in the acceleration of orofacial gestures. The "Probabilistic Reduction Hypothesis" (Gregory, Raymond, Bell, Fosler-Lussier, & Jurafsky, 1999) precisely describes the articulatory reduction of the predictable (emulated) linguistic items in speech production.

4.1.2. *Semantic bleaching and proceduralization of conceptual representations*

Semantic bleaching has been attributed to habituation processes: the organism ceases to exhibit the same response strength to frequently occurring stimuli (Haiman, 1994). A strong neural candidate is the attenuation of the actual sensory consequences as compared with the CB predictions (Blakemore, Wolpert, & Frith, 2000). Gating of sensory information heavily involves the BG (see **section 4.2**). "Shallow processing" (Barton & Sanford, 1993) and "good-enough representations" (Ferreira, Bailey, & Ferraro, 2002) capture aspects of minimized attentional costs in semantic processing that a routinization-induced low KG modality may achieve. To the extent that processing efficiency increases, semantic representations of words and constructions are underspecified, and ultimately bleached.

However, semantic bleaching expands the contexts of occurrence of linguistic items, inviting non-conventional inferences (Heine *et al*., 1991). While such higher cognitive inferential processes should heavily involve the cognitively demanding exploration of the temporoparietal cortex (the putative conceptual repository), grammaticalization does not occur but with the "proceduralization" of the conceptual representations that such non-conventionalized inferences invoke. Procedural encoding provides the "necessary processing constraint on the interpretation of an associated conceptual representation" (Nicolle, 1998, p.23). Characteristically, while it "performs the same role in constraining or guiding the interpretation of the utterance that an increase in the number of lexical items can have" (LaPolla, 2003, p.135), procedural encoding is "automatically recovered (in addition to being merely activated on decoding)" (Nicolle, 1998, p.23).

Proceduralization reflects KG minimization in semantic processing **(figure 3)**: The "cognitive cerebellum" may (redundantly) emulate the subconscious "mental background", e.g., the rules of a game, constraining the conscious, cortical "mental foreground", e.g., planning for a winning strategy (Thach, 1998): A CNMC might connect to the cerebral loop as a reliable copy of the thought model in the temporoparietal areas, with the thought process being alternatively conducted by the frontal areas acting on the CNMC rather than on the temporoparietal areas, adaptively avoiding the conscious effort needed for the exploration of cortical loci (Ito, 2000).

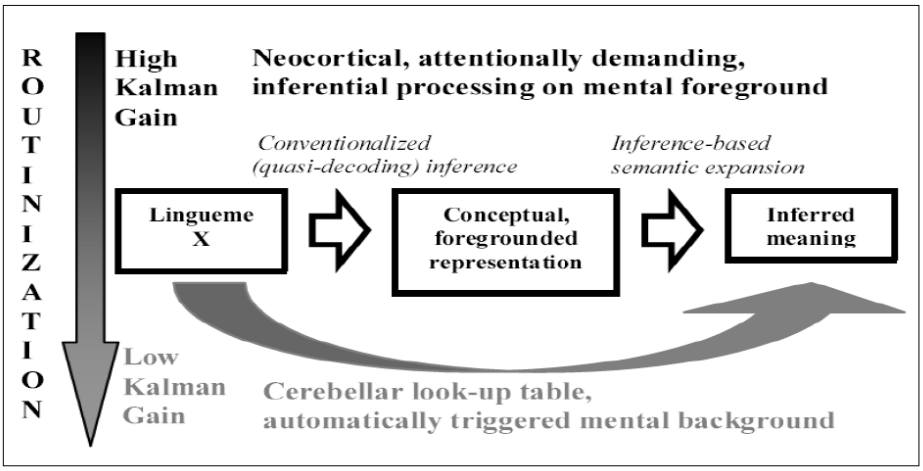


Figure 3. Routinization-induced proceduralization of conceptual encoding meets the "cerebellarization" of cognitive repertoires.

4.2. *Striatal Regulation of Cortical Unification Operations*

The fuzziness of syntactic categoriality, emphasized by grammaticalization theorists (e.g., Givón, 1979), has recently attracted researchers from computational/psycholinguistic probabilistic modeling (Zuraw, 2003 for a review), encouraging the definition of categoriality on the basis of the particular constructions that each item occurs in. In Pulvermüller's (2002) neuronal syntax, lexical categories are defined by the set of the very complements lexical categories require, i.e., by their "sequence regularities" (*ibid.*). An efficient parser thus gates candidates for unification based on the context-sensitive inhibitory strengths of their connections to their competitors; this is directly reflected in Vosse and Kempen's (2000) model, and is implementable by Pulvermüller's (2002) "striatal regulation of cortical activity".

Characteristically, Walenski, Mostofsky, & Ullman (2007) report particularly speeded processing of procedural (both linguistic and non-linguistic) knowledge for Tourette's syndrome subjects, attributing it to their BG

abnormalities in the inhibition of frontal cortical activity. Grossman, Lee, Morris, Stern, and Hurtig (2002) found a correlation between sentence comprehension and Stroop task performance in Parkinsonians, while Hochstadt, Nakano, Lieberman, and Friedman (2006) attributed their compromised capacity of parsing relative clauses to "deficits in cognitive set-switching" or "underlying inhibitory processes". Inhibition and reinforcement underlie probabilistic representation: BG patients exhibit deficient probabilistic category learning (Knowlton *et al*., 1996). In its acquisition phase, striatal activation is involved for normal individuals (Poldrack, Prabhakaran, Seger, & Gabrieli, 1999).

Thus, grammaticalization-induced decategorialization, becoming manifest with alterations in E- language distributional patterns, is efficiently monitored by BG reinforcement learning, via the dopamine-mediated regulation of the inhibitory strengths among syntactic variants that compete for unification with a particular linguistic item **(figure 4)**.

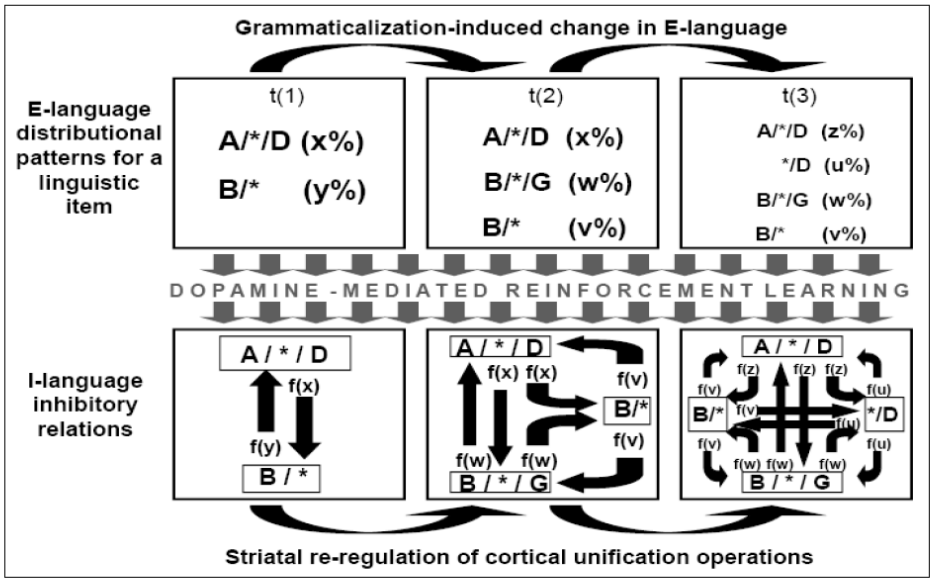


Figure 4. **t(1)-t(2):** a member of category G frequently co-occurs with the sequence B/*/, which triggers the strengthening of its probabilistic representation in the frontostriatal circuit, and thus the strength of the inhibitory signals sent to the competing alternatives: a gradual "obligatorification" (Lehmann, 1995). **t(2)-(3):** a member of category A initiating the sequence A/*/D becomes optional, i.e., outcompeted in the winner-takes-all BG selections for cortical linguistic unifications.

5. Conclusion

I proposed two fundamental neurolinguistic mechanisms grounding grammaticalization operations: a) the cerebellar-induced Kalman gain reduction in linguistic processing, and b) the basal ganglionic adaptive regulation of cortical unification operations.

Acknowledgments

My special thanks to Prof. Jim Hurford, Dr. Thomas Bak, Prof. Mike Paulin, Dr. Holly Branigan, Dr. Patrick Sturt, Dr. Andrew Smith, and Dr. Anna Parker. I gratefully acknowledge the Pavlos and Elissavet Papagiannopoulou Foundation for the financial support for my PhD studies.

References

Ackermann, H., & Hertrich, I. (2000). The contribution of the cerebellum to speech processing. *Journal of Neurolinguistics*, *13*, 95-116.

Barton, S.B., & Sanford, A.J. (1993). A case-study of anomaly detection: shallow semantic processing and cohesion establishment. *Memory and Cognition*, *21*, 477-487.

Blakemore, S.-J., Wolpert, D., & Frith, C. (2000). Why can't you tickle yourself? *NeuroReport*, *11*, 11-16.

Bybee, J. L. (1998) A functionalist approach to grammar and its evolution. *Evolution of Communication*, *2*, 249-278.

Bybee, J. L., & Thompson, S. (2000). Three frequency effects in syntax. *Berkeley Linguistics Society*, *23*, 65-85.

Doya, K. (1999). What are the computations of the cerebellum, the basal ganglia and the cerebral cortex? *Neural Networks*, *12*, 961-974.

Fabbro, F., Tavano, A., Corti, S., Bresolin, N., De Fabritiis, P., & Borgatti, R. (2004), Long-term neuropsychological deficits after cerebellar infarctions in two young adult twins. *Neuropsychologia*, *42*, 536-545.

Ferreira, F., Bailey, K. G. D., & Ferraro, V. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science*, *11*, 11–15.

Fiez, J.A., Petersen, S.E., Cheney, M.K., & Raichle, M.E. (1992). Impaired non-motor learning and error detection associated with cerebellar damage. *Brain*, *115*, 155-178.

Givón, T. (1979). *On Understanding Grammar*. New York: Academic Press.

Graybiel A. M. (2000). The basal ganglia. *Current Biology*, *10*, 509–11.

Gregory, M. L., Raymond, W. D., Bell, A., Fosler-Lussier, E., & Jurafsky, D. (1999). The effects of collocational strength and contextual predictability in lexical production. *Proceedings of the Chicago Linguistic Society*, *99*, 151–166.

Grossman, M., Lee, C., Morris, J., Stern, M.B., & Hurtig, H.I. (2002). Assessing resource demands during sentence processing in Parkinson's disease. *Brain and Language*, *80*, 603-616.

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

Haiman, J. (1994). Ritualization and the development of language. In William Pagliuca (Ed.), *Perspectives on Grammaticalization*, (pp. 3-28). Amsterdam: John Benjamins.

Heine, B., & Reh, M. (1984). *Grammaticalization and Reanalysis in African Languages.* Hamburg: Helmut Buske.

Heine, B., Claudi, U., & Hünnemeyer, F. (1991). From cognition to grammar. Evidence from African languages. In E. C. Traugott and B. Heine (Eds.), *Approaches to grammaticalization*, Vol. 2 (pp. 149-187). Amsterdam: Benjamins.

Hochstadt, J., Nakano, H., Lieberman, P., & Friedman, J. (2006). The roles of sequencing and verbal working memory in sentence comprehension deficits in Parkinson's disease. *Brain and Language*, *97*, 243-25.

Ito, M. (1984). *The Cerebellum and Neural Control*. New York: Raven Press.
Ito, M. (2000). Neural control of cognition and language. In A. Marantz, Y. Miyashita, and W. O'Neil (Eds.), *Image, language, brain* (pp. 149-162). Cambridge, MA: MIT Press.
Knowlton B.J., Squire, L.R., Paulsen, J.S., Swerdlow, N.R., Swenson, M., & Butters, N. (1996). Dissociations within non-declarative memory in Huntington's disease. *Neuropsychology*, *10*, 538–48.
LaForce, R., & Doyon, J. (2001). Distinct contribution of the striatum and the cerebellum to motor learning. *Brain and Cognition*, *45*, 189-250.
LaPolla, R. J. (2003). Why languages differ: Variation in the conventionalization of constraints on inference. In D. Bradley, R.J. LaPolla, B. Michailovsky, and G. Thurgood (Eds.), *Language Variation: Papers on Variation and Change in the Sinosphere and in the Indosphere in Honour of James A. Matisoff* (pp. 113-144). Canberra: Pacific Linguistics, Australian National University.
Lehmann, C. (1995). *Thoughts on grammaticalization*. Munich: Lincom Europa. (First published as *akup* 48, Institut für Sprachwissenschaft, Universität zu Köln, 1982).
Nicolle, S. (1998). A relevance theory perspective on grammaticalization. *Cognitive Linguistics*, *9(1)*, 1-35.
Paulin, M.G. (1989). A Kalman filter theory of the cerebellum. In M.A. Arbib and S.-I. Amari (Eds.), *Dynamic Interactions in Neural Networks: Models and Data* (pp. 239-259). New York: Springer.
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)*, 105-110.
Poldrack, R. A., Prabhakaran, V., Seger, C. A., & Gabrieli, J. D. (1999). Striatal activation during acquisition of a cognitive skill. *Neuropsychology, 13(4),* 564–574.
Pulvermüller, F. (2002). *The neuroscience of language: On brain circuits of words and serial order*. Cambridge: Cambridge University Press.
Stowe, L.A., Paans, A.M.J., Wijers, A.A., & Zwarts, F. (2004). Activations of "motor" and other non-language structures during sentence comprehension. *Brain and Language*, *89*, 290-299.
Thach, W.T. (1998). What is the role of the cerebellum in motor learning and cognition? *Trends in Cognitive Sciences, 2(9)*, 331-337.
Thach, W. T., Mink, J. W., Goodkin, H. P., & Keating, J. G. (2000). Combining versus gating motor programs: Differential roles for cerebellum and basal ganglia. In M. S. Gazzaniga (Ed.), *Cognitive neuroscience: A reader* (pp. 366–375). Oxford: Blackwell.
Vosse, T., & Kempen, G. (2000). Syntactic structure assembly in human parsing: a computational model based on competitive inhibition and a lexicalist grammar. *Cognition*, *75*, 105-143.
Walenski, M., Mostofsky, S.H., & Ullman, M.T. (2007). Speeded processing of grammar and tool knowledge in Tourette's syndrome. *Neuropsychologia, 45*, 2447-2460.
Wolpert, D.M., Miall, R.C., & Kawato, M. (1998). Internal models in the cerebellum *Trends in Cognitive Sciences*, *2(9)*, 338-347.
Zuraw, K. (2003). Probability in Language Change. In R. Bod, J. Hay and S. Jannedy (Eds.), *Probabilistic linguistics*. (pp. 139-176). Cambridge, MA: MIT Press.

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.

Abstract

Despite the growing literature of empirical findings in support of the involvement of the cerebellum in language processing, the identity of the underlying linguistic computations per se has attracted little, if any, attention at all. Emphasizing language comprehension here, a neurolinguistic model of the posterolateral cerebellum is proposed, grounded on a recently formulated psycholinguistic processor, capturing aspects of rapid, predictive, noise-resistant, and covertly imitative language comprehension processes.

1. Introduction

The lateral cerebellum (henceforth CB) exhibits a pronounced reciprocal expansion with frontal cortical areas (Dow, 1942), and a striking volume increase in hominoids as opposed to monkeys (e.g. MacLeod et al., 2003; Whiting & Barton, 2003). In particular, its reciprocal connectivity with Broca's and Wernicke's area as well as with the dorso-lateral prefrontal cortex (Leiner et al., 1991; Schmahmann & Pandya, 1997; Middleton & Strick, 1998), along with its cytoarchitectural homogeneity, suggestive of unitary CB computations (e.g. Bloedel, 1992; Schmahmann, 1997; Wolpert et al., 1998), encourage the formulation of a neurolinguistic model that makes reference to psycholinguistic processes pertaining to such CB neurocomputations. However, despite the psychopathological evidence for a *lateralized linguistic cerebellum* (Mariën et al., 2001), little if anything has been said about the involvement of such computations per se in language processing, with research on the CB linguistic psychopathology predominantly constrained to the examination of *frontal-like* syndromes of language disorders that CB deficits might induce, either because of the functional deafferentization of the cortical language-related loci, or because of the disruption of the CB modulatory role, as in verbal working memory (Silveri, 1994; Molinari et al., 1997). The present paper provides a synthesis of psycholinguistic models of language comprehension processes with work on the computations of the posterolateral CB.

2. State estimation in language comprehension

Wilson & Knoblich (2005) had reviewed the involvement of covert imitative production mechanisms in the efficient and noise-resistant perception of conspecifics in contexts of increased predictability (*low Kalman gain*). Based on their observations, Pickering and Garrod (2007) propose that production holds a causal role in the efficient perception of noisy/ambiguous linguistic input and in online prediction, and construct a psycholinguistic processor that uses a production-based language *emulator* (Grush, 2004), which is *controlled by feedback from a Kalman filter* (Pickering & Garrod, 2007, p. 108), weighing predictions against analysis of the input at each step. In cases where the prediction is strong and the input noisy, the internal model the language production system provides exerts strong top-down influence to the *input analysis system.* Among others, the processor captures cases of on-line prediction in multi-level language processing (see Pickering & Garrod (2007) for a brief review), noise resistance in speech perception (e.g. phoneme restoration phenomena (Warren, 1970), where listeners are unable to detect a phone that is occluded by physically similar noise, based on expectations generated in word or *a fortiori* sentential contexts), cases of *shallow processing* (Sanford & Sturt, 2002) in sentence comprehension (as manifested, for instance, in Wason and Reich's (1979) *verbal illusions,* e.g. *No head injury is too trivial to be ignored*), as well as cases of involvement of speech production mechanisms in speech perception (e.g. Watkins et al., 2003).

3. The model

Considerations of the neural instantiation of such psychological processes have primarily involved the mirror neuron literature: Wilson & Knoblich (2005) explicitly identify mirror neurons as the neural foundation of the covert imitative involvement of the production system in the perception of conspecifics in contexts of increased predictability. In the same vein, Iakoboni (2005) and Hurley (2008) use the concepts of internal models and emulation to capture the function of mirror neurons in higher cognitive processes. However, what has been ignored in such considerations is the CB (Miall, 2003): the *Cerebellar CorticoNuclear MicroComplex* (henceforth CCNMC), i.e., the fundamental CB functional unit, provides the ground of *internal* (Ito, 1984), or *forward models* (Kawato et al., 1987)[1], or *emulators* (Grush, 2004), or, similarly, the neural analogue of the Kalman Filter (Paulin, 1989, 1997).

A CCNMC consists of a microzone, i.e., a small area of the CB cortex, and a small number of nuclear cells. Simply put, the microzone receives two kinds of input, mossy fibers and climbing fibers, and the output is carried by the deep CB nuclear cells. The set of mossy fiber inputs is transformed by the granule cells whose axons form the parallel fibers. The axons of the Purkinje cells, the only output cells of the microzone, are sent to a small group of vestibular or CB nuclear neurons. Long Term Depression (LTD) occurs at parallel fiber-to-Purkinje cell synapses after conjunctive activation of these synapses together with climbing fiber (transmitting the error signals)-to-Purkinje cell synapses (Ito et al., 1982).

The *NeoCerebellar Kalman Filter Linguistic Processor* (henceforth NCBKFLP; see Figure 1) is proposed here to provide the neurofunctional grounding of Pickering and Garrod's (2007) psycholinguistic processor. Closely following the functional principles of Ito's CCNMCs, the linguistic internal models of the NCBKFLP would learn on the basis of the discrepancies between the predicted state and the actual state that the cortically instantiated *input analysis system* (*ibid.*) enters: CB predictions transmitted by the ventrolateral neodentate nucleus via the phylogenetically newer parvocellular part of the red nucleus and the ventrolateral nucleus of the thalamus to the mirror neurons of Broca's area are trained by the error signals transmitted via the climbing fibers, with long-term depression (LTD) occurring at the synapses of Purkinje cells-parallel fibers of the posterolateral CB cortex. The inferior olivary nuclei receive signals both directly from the dentate nucleus and indirectly (see Ito, 2008; Schmahmann, 2001 for references) from prefrontal association areas via the zona incerta. In that way, the predictions of the CB linguistic internal model can be compared with the actual output of the cortical input analysis system. An efference copy of the state of the cortical input analysis system is transmitted by the cerebro-ponto-cerebellar pathway (see Schmahmann & Pandya, 1997), with mossy-parallel fibers reaching the Purkinje cells of the posterolateral CB. A CCNMC might thus connect to the cerebral loop as a reliable copy of the language model in Wernicke's area, with the language process being alter-

1 In the motor domain, forward models predict the sensory consequences of movements from an efference copy of issued motor commands. Inverse models compute necessary feedforward motor commands from desired movement information. Both kinds of internal models are assumed to be predominantly located in the cerebellum (Kawato, 1999).

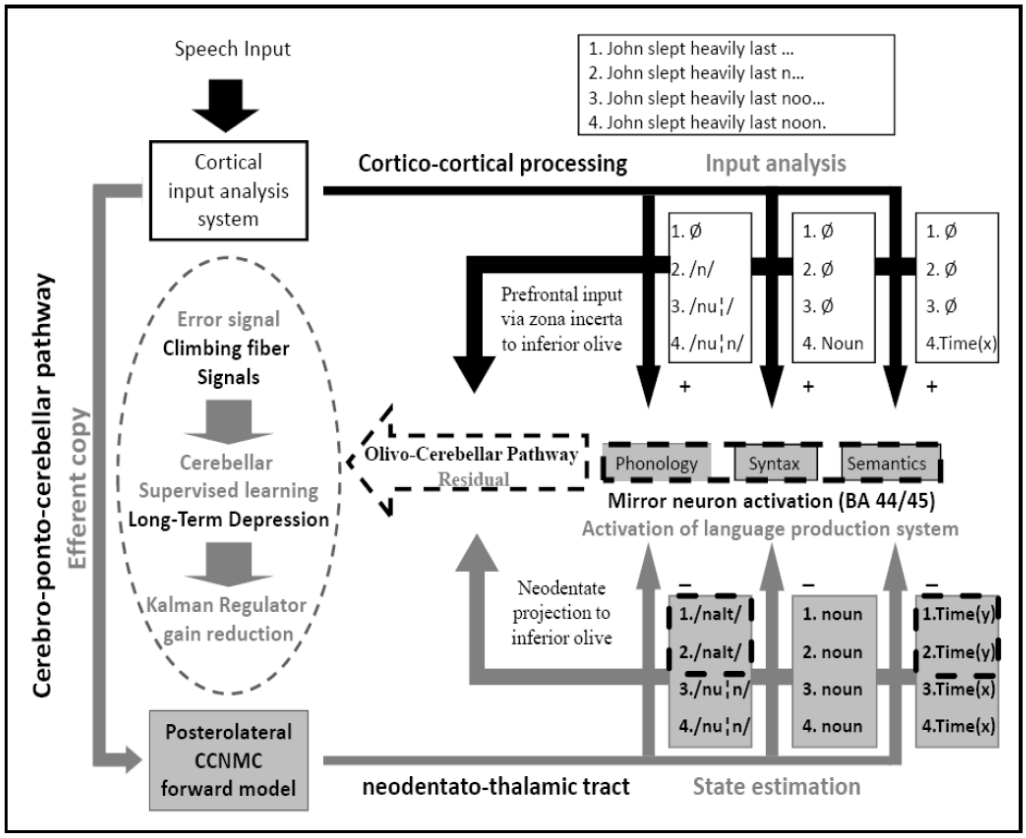


Figure 1: The Neo-Cerebellar Kalman Filter Linguistic Processor provides the neural grounding of Pickering & Garrod's (2007) processor: Mossy fiber inputs, originating at the pontine nuclei, convey, via the parallel fibers, the state of the cortical input analysis system to Purkinje cells at the (right) posterolateral CB cortex. Purkinje cells via their inhibitory connection to the ventrolateral part of the dentate nucleus, via the parvocellular red nucleus and the ventrolateral thalamus, signal the expected state to the mirror neurons of Broca's area, triggering the covert involvement of the language production system. Discrepancies between the CB predictions and the actual sensory input are conveyed as error signals back to the posterolateral CB cortex, via the climbing fibers that originate at the inferior olivary nuclei, which receive indirect connections from the cortical language related areas via the zona incerta and direct projections from the dentate nucleus . Long Term Depression occurs at the parallel fiber to Purkinje cell synapses of the posterolateral cerebellar cortex after conjunctive activation of these synapses together with climbing fiber-to-Purkinje cell synapses.

natively conducted by Broca's area acting on the CCNMC rather than on Wernicke's, adaptively avoiding the conscious effort needed for the exploration of cortical loci (Ito, 2000, 2008).

Evidence for CB error signalling and error-driven learning in the linguistic domain has already been provided: Fiez et al.'s (1992) CB patient showed deficient supervised learning in a concurrent discrimination word learning task, and likewise failed to detect instances of his erroneous performance in word generation tasks. In an fMRI study of processing syntactically ambiguous sentences, significant CB activation was found at the point of reanalysis of the sentence (Stowe et al., 2004), suggestive of the CB involvement in error signalling (Jueptner & Weiller, 1998). In the same vein, the gain of a disrupted Kalman regulator cannot be lowered on the basis of the residuals: lack of practice-induced facilitation in linguistic tasks is characteristic in CB patients, who poorly reduce their reaction times and accuracy rates across blocks, as opposed to normal controls in a variety of word generation tasks (Fiez et al., 1992; Gebhart et al., 2002). These findings reflect a more domain-general pattern of lack of performance optimization in CB deficits, most recently demonstrated in Ferrucci et al. (2008), where CB Transcranial Direct Current Stimulation specifically impaired the practice-dependent proficiency increase in verbal working memory, while TDCS over the prefrontal cortex induced an immediate change in the WM task but left the practice-dependent proficiency unchanged.

Some first suggestive evidence for the involvement of the CB in the restoration of noisy linguistic percepts comes from studies in dyslexia accomodatable within the framework of the *Cerebellar Deficit Hypothesis* (Nicolson et al., 2001): the pronounced difficulties of dyslexic subjects in perceiving speech in noisy contexts (Sperling et al., 2005) has been attributed to deficient CB internal models in speech processing (Ito, 2008). The CB predictions matching the input processed by the cortical input analysis system induce *cancellation of reafference* (Nelson & Paulin, 1995; Blakemore et al., 1998), which, at the phonological level, would explain why white noise in an occluded fricative is harder to detect in contexts of highly frequent words (strong top-down influence) as opposed to infrequent ones (Samuel, 1981).

Not noticing the evidence (Sanford & Sturt, 2002) in the cases of *verbal illusions* (Wason & Reich, 1979), *pragmatic normalization* (Fillenbaum, 1974), and *shallow processing* phenomena seem to involve the same CB function at the sentential level, where an incorrect mental model is constructed (Garnham & Oakhill, 1996), bypassing the processes of the syntax-semantics interface and providing an interpretation based on the habitual arrangement of the semantic components of the parsed sentence. Internal models of the posterolateral CB have been considered to acquire the dynamics of the manipulation of a particular mental model stored in the temporoparietal cortex as the control object, and thus facilitate processing below the level of awareness (e.g. Ito, 2000). The CB may thus pre-emptively output an arrangement of the semantic components (e.g, `dog´(x), man´(y), bite´(AGENT, PATIENT)`) of the input sentence (e.g. the often pragmatically normalized sentence *The man bit the dog*) based on their predicted, statistically prominent arrangement `(dog´(x), man´(y), bite´(x,y))`, compensating for the often noisy/ambiguous linguistic input, and bypassing the slower cortical com-

putations of the syntax/semantics interface that would guarantee the correct, yet slower and infrequent interpretation (bite´(y,x)). Suggestively, the CB has already been shown to be involved in non linguistic illusions of state estimation, as the *size-weight illusion* (Hubbard & Ramachandran, 2004), where CB patients showed a minimized effect of the illusion to different degrees as opposed to normal control subjects.

In the model thus proposed, the connectivity between mirror neurons in Broca's area and the right posterolateral CB guarantees a functional cooperation of imitation and emulation mechanisms, respectively. While the properties of the connectivity of BA 44/45 mirror neurons with the right posterolateral CB have not yet been researched, it has suggestively been held that CB afferents via the dentatothalamic tract to the F5, i.e., Broca's homologue in monkeys (Petrides & Pandya, 1994) might have input-output relations with the mirror neurons in that area representing the *desired state* in tool-use learning (Imamizu et al., 2003). In the same vein, the estimates of the right posterolateral CB emulator (in cases of a low Kalman regulator gain) would be transmitted to the mirror neurons of Broca's area, triggering the covert employment of production mechanisms in language comprehension.

4. Cortical rules and cerebellar habits

Efferent copies of sentential inputs of the form of [N(oun), V(erb), N(oun)] can be transmitted from the cortical input analysis system to the Purkinje cells of the posterolateral CB cortex via the mossy fiber afferents (the cerebro-ponto-cerebellar pathway). Indeed, the length of the parallel fibers is sufficient to synaptically connect many microzones (e.g. Mugnaini, 1983) and transmit composite contexts to the microcomplexes, while the mediating granule cells ensure that each parallel fiber carries a combination of activity on several mossy fibers, with parallel fibers thus conveying combined contextual representation to Purkinje cells. Combinations of such elementary linguistic categories can be trained in the same fashion that elementary moves are trained into compounds in the CB (Thach et al., 1992). This linguistic internal model, via routinization, can reliably copy the dynamics of the thematic role assignment process, rapidly providing to such input sequences the output of the form [Agent = N1, Action = V, Theme = N2], with the neodentate pre-emptively transmitting this information back to the language-related cortical loci, ensuring the ease and speed of processing active sentences (Forster & Olbrei, 1974).

The cortical neurolinguistic chauvinism take on subcortical structures has often marginalized findings of subcortical linguistic deficits not reflecting *frontal-like* patterns. The most prominent example is Pickett (1998), the first report in the literature of sentence comprehension deficits attributable to CB pathology, where normal control subjects made errors in the interpretation of 4% of the active sentences and 10% of the passive sentences, whereas CB patients, surprisingly, made errors on 11% of both active and passive sentences, showing no effect of Voice manipulation, with the lateral CB patients of the overall group strikingly performing better than normal control subjects. Such performance patterns, interpretable as *mixed results* (Justus, 2004) from the *fron-*

tal-like perspective, are even more clearly contrasted by the ones of Broca's agrammatic aphasics in comprehending passive sentences (e.g. Saffran et al., 1998), who exhibit an exaggerated degree of such heuristics-induced (strategic or not) misinterpretations.

According to Townsend & Bever's "Late Assignment of Syntax Theory" (LAST, 2001), sentence comprehension employs both algorithmic (the actual *syntax*) and faster, pre-emptive *heuristic, associative* computations (the *pseudosyntax*), the latter utilizing statistically valid information to elicit an initial meaning/form hypothesis, based on which the former will construct a detailed syntactic analysis. The NCBKFLP captures such cases of two-stage serial models of sentence comprehension, with the *habits/rules* divide being instantiated in the slower cortico-cortical versus the faster (Doya, 1999) corticocerebellar modality of language processing, respectively: the *emulational* function of the CB advocates in favour of its involvement in such *pseudosyntactic* operations, as opposed to the actual, algorithmic processes or the very representations of the syntactic templates per se. Characteristically, the notion of *neurofunctional redundancy* has been invoked for the CB (emulated) linguistic representations, as CB aphasia is significantly milder than classical aphasic syndromes, owing to maximal prefrontal compensation (Fabbro et al., 2004). The CB is also fundamental in associative learning in both motor and higher cognitive tasks (Drepper et al., 1999), also forming in many comparative neurocognitive models a fundamental locus of habits (Mishkin et al., 1984), the domain generality of which, supported by its cytoarchitectural homogeneity and its reciprocal connectivity with the language-related cortical loci makes it promisingly inclusive of syntax and semantic processing habits. Thus, with the abolishment of the NVN pseudoparse, i.e., the heuristic of the highly entrenched theta-role assignment of [Agent = N1, Action = V, Theme = N2], the overall *cost for the passive structure* (Ferreira, 2003) now pertains equally to the active voice sentences as well. This provides a concrete neurolinguistic basis for Pickett's (1998, p. 103) own suggestion that *the linguistically impaired cerebellar subjects… can perform the full range of linguistic processing, but their mental shortcuts have been disrupted, so that processing 'easy' sentences requires more conscious effort and is performed less automatically.* The *weak coupling* of the internal model with the copied target system (Paulin, 1997) in language processing-state estimation thus translates into the *good-enough* representations constructed by hearers/readers in sentence comprehension (Ferreira et al., 2002) and the subsequent misinterpretations induced.

Townsend and Bever (2001), assuming a different perspective on the neural instantiation of LAST, propose that pseudosyntactic operations are instantiated in Broca's area, holding that the NVN pseudoparse seemingly found in agrammatic aphasics is but a non-linguistic strategy. However, this immediately makes the proposal that the NVN heuristic is supported by Broca's area run into unfalsifiability issues. Moreover, Broca's area is best construable as a part of the *procedural* cortico-striatal and cortico-cerebellar circuit (Ullman, 2004), so, certain pseudosyntactic operations are conceivably CB-dependent. Finally, Townsend & Bever's proposal (2001) heavily relies on the findings of Linebarger et al. (1983) on the preservation of grammatical knowledge in Broca's aphasics, against which, however, Grodzinsky & Finkel (1998) have shown that agrammatic aphasics exhibit severe deficits in grammaticality judgment tasks involving

structures with dependencies between traces and their antecedents, thus showing (regardless of the syntactic framework employed) that the agrammatic deficit is structure-, but not task-dependent. The CB thus appears to be at least one neuroanatomical structure supporting pseudosyntactic operations, and at most the only locus undertaking such linguistic operations.

5. The NCBKFLP in grammaticalization

Grammaticalization is definable as the diachronic process that leads lexical items to grammatical ones inside constructional morphosyntactic contexts. In Argyropoulos (2008), it was proposed that the neocerebellum and the basal ganglia might provide the neural foundation for the domain-general automatization phenomena that constitute the cognitive core of grammaticalization operations (Givón, 1979, and especially 1989; Bybee, 1998; Lehmann, 2004), and that current neurolinguistic research on those structures allows the first steps towards the articulation of a neurolinguistics of grammaticalization.

Given that the products of grammaticalization operations are introduced as instances of optimized linguistic performance at the intra-generational level of language transmission (Haspelmath, 1998), it is important to consider the neurocognitive mechanisms underlying the optimization of repeatedly processed linguistic repertoires, rather than the storage of linguistic representations per se: the significance of the CB in the optimization of thought and language processing, as first proposed by Leiner et al. (1986), promotes this structure as a promising candidate, *a fortiori* given its fundamental role in the practice-induced shift from a controlled (prefrontal and anterior cingulate cortex) to an automated (Sylvian-insular) language processing circuit in linguistic tasks (see Fiez & Raichle, 1997), i.e., the neurocognitive shift suggested to underlie the realization of grammaticalization processes (Givón, 1989).

The very construal of *pseudosyntax* as the process stage of *syntactic habits* (Townsend & Bever, 2001) presupposes the involvement of adaptive demotions from stage 2 to stage 1 operations that routinization of language processing would induce: characteristically, Kempson & Cann (2007) account for the procliticization (see [1], taken from Bouzouita, 2002) of object pronouns in the transition from Medieval to Renaissance Spanish as *a natural subsequent step of routinization*, involving the calling up of *actions associated with the verb together with those associated with the clitic with a single lexical look-up mechanism*:

[1] mas los rompan luego
but CL break.3PL afterwards
'but break them afterwards'

CB supervised learning provides such a *look-up table* or *shortcut-circuit* for mappings originally developed by *the time-consuming cortico-cortical processing* (Doya, 1999, p. 970). Evidence for the involvement of the right posterolateral CB in language pro-

cessing in a constrained search space has been provided in imaging and behavioural studies: increased CB activation was found in the *FEW* condition in a stem completion task, i.e., where the candidate completions of the stem were few in number (e.g. *PSA-*), as opposed to increased frontostriatal activation in the *MANY* condition (e.g. *STA-*), interpreted as involving the selection of a particular response among competing alternatives (Desmond et al., 1998). Gebhart et al. (2002) lend similar emphasis to the size of the search space as a factor affecting the performance of their right posterolateral CB patients, as performance is spared in the subordinate category generation (e.g. stimulus: *MONEY*, response: *dollar, pound, euro,* etc.) task, where many correct word responses are available, yet is poor in the antonym (e.g. stimulus: *KIND*, response: *rude*) and verb generation (e.g. stimulus: *CHEF*, response: *cook*) tasks, where few correct word responses are available.

The *Linear Fusion Hypothesis*, predicting that *items that are used together fuse together* (Bybee, 2002), applies widely in grammaticalization operations, with affixation, agglutination, and merger as characteristic cases of universal *syntagmatic coalescence* (Lehmann, 1995) phenomena. Independently of whether chunking in motor/cognitive behaviour is supported by CB state estimation (Paulin, 1989) or by a CB composing function (Thach et al., 1992), the NCBKFLP directly accounts for the automated chunks that hearers process (Bybee, 2002) on the basis of the covert generation of an anticipated co-occurring item B in response to a perceived item A. The impaired anticipatory planning found early in CB patients (Leiner et al., 1987), as well as the severe impairments found in cognitive associative learning (Drepper et al., 1999) advocate in favour of such a kind of CB involvement.

Finally, the core, semantic change in grammaticalization processes (as opposed to the secondary aspect of the formal and further semantic reductions that accompany such shift) has been accounted for by grammaticalization researchers in terms of a demotion from a cognitive *foreground* to a cognitive *background* in routinized discourse processing (Boyer and Harder, forthcoming; Lehmann, 2004). Given the above considerations then, the fact that neuroscientists have used identical terms and notions to explicate the significance of the CB in the automatization of thought processing (Thach, 1998) is suggestive of a perspective of a fruitful synthesis that historical linguistics may reach with recent developments in the field of the neuropsychology of language.

6. Conclusion

The introduction of computations of state estimation in psycholinguistic models provides the cornerstone for the articulation of concrete neurolinguistic models of the posterolateral cerebellum beyond its commonly held strictly modulatory role, along with a set of falsifiable experimental hypotheses to be assessed in forthcoming work by the author.

7. Acknowledgements

My special thanks to Prof. Merle Horne and Mikael Roll for their support, as well as to Dr. Thomas Bak, Dr. Kasper Boye, Dr. Ronnie Cann, Prof. Fernanda Ferreira, Prof. Talmy Givón, Prof. Peter Harder, Prof. Martin Haspelmath, Prof. James Hurford, Prof. Christian Lehmann, Prof. Chris Miall, Prof. Martin Pickering, Dr. Patrick Sturt, and Prof. David Townsend for the encouragement and their helpful comments – I hold exclusive responsibility for any mistakes here. I gratefully acknowledge the Pavlos and Elissavet Papagiannopoulou Foundation for the financial support.

References

Argyropoulos, G. P. (2008). The subcortical foundations of grammaticalization. In A. D. M. Smith, K. Smith, & R. F. i Cancho (Eds.), *The Evolution of Language: Proceedings of the 7th International Conference on the Evolution of Language* (pp. 10–17). Singapore: World Scientific Press.

Barton, S.B., & Sanford, A.J. (1993). A case-study of anomaly detection: Shallow semantic processing and cohesion establishment. *Memory and Cognition*, *21*, 477–487.

Blakemore, S.-J., Wolpert, D. M., & Frith, C. D. (1998). Central cancellation of self-produced tickle sensation. *Nature Neuroscience, 1*(7), 635–640.

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

Bouzouita, M. (2002). Clitic placement in Old and Modern Spanish: A dynamic Account. MSc. Dissertation. King's College London.

Boye, K., & Harder, P. (forthcoming). Evidentiality: Linguistic categories and grammaticalization. *Functions of Language.*

Bybee, J. L. (1998). A functionalist approach to grammar and its evolution. *Evolution of Communication*, *2*, 249–278.

Bybee, J. L. (2002). Sequentiality as the basis of constituent structure. In T. Givón & B. Malle (Eds.) The Evolution of Language out of Pre-Language (pp. 107–132). Amsterdam: John Benjamins.

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,* 368–376.

Dow, R.S. (1942). The evolution and anatomy of the cerebellum. *Biological Reviews of the Cambridge Philosophical Society, 17,* 179–220.

Doya, K. (1999). What are the computations of the cerebellum, the basal ganglia and the cerebral cortex? *Neural Networks, 12,* 961–974.

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.

Fabbro, F., Tavano, A., Corti, S., Bresolin, N., De Fabritiis, P., and Borgatti, R. (2004), Long-term neuropsychological deficits after cerebellar infarctions in two young adult twins. *Neuropsychologia, 42,* 536–545.

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

Ferreira, F., Bailey, K. G. D., & Ferraro, V. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science, 11*, 11–15.

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), 1687–1697.

Fiez, J. A., Petersen, S. E., Cheney, M. K., & Raichle, M. E. (1992). Impaired non-motor learning and error detection associated with cerebellar damage. *Brain, 115,* 155–178.

Fiez, J. A., & Raichle, M. (1997). Linguistic processing. In J. D. Schmahmann (Ed.), *The Cerebellum and Cognition* (pp. 233–54). San Diego: Academic Press.

Fillenbaum, S. (1974). Pragmatic normalization: Further results for come conjunctive and disjunctive sentences. *Journal of Experimental Psychology, 102*, 574–578.

Forster, K. I., & Olbrei, I. (1973). Semantic heuristics and syntactic analysis. *Cognition*, *2*, 319–347.

Garnham, A., & Oakhill, J. V. (1996). The mental models theory of language comprehension. In B. K. Britton & A. C. Graesser (Eds.), Models of understanding text (pp. 313–339). Hillsdale, NJ: Lawrence Erlbaum Associates.

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,* 318–333.

Givón, T. (1989). *Mind, code and context: Essays in pragmatics.* New York: Academic Press.

Givón, T. (1979). *On Understanding Grammar*. New York: Academic Press.

Grodzinsky, Y., & Finkel, L. (1998). The neurology of Empty Categories. Aphasics' Failure to detect ungrammaticality. *Journal of Cognitive Neuroscience, 10*(2), 281–292.

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

Haspelmath, M. (1998). Does grammaticalization need reanalysis? *Studies in Language, 22,* 315–51.

Hubbard, E. M., & Ramachandran, V.S. (2004). The Size-Weight Illusion, Emulation, and the Cerebellum. *Behavioral and Brain Sciences, 27,* 407–408.

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.

Iacoboni M. (2005). Understanding others: Imitation, language, empathy. In S. Hurley & N. Chater (Eds.) *Perspectives on Imitation: From Cognitive Neuroscience to Social Science, Volume 1: Mechanisms of Imitation and Imitation in Animals* (pp. 77–99). Cambridge, MA: MIT Press.

Imamizu, H., Kuroda, T., Miyauchi, S., Yoshioka, T., & Kawato, M. (2003). Modular organization of internal models of tools in the human cerebellum. *Proceedings of the National Academy of Sciences of the United States of America, 100*(9), 5361–5466.

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

Ito, M. (2000). Neural control of cognition and language. In A. Marantz, Y. Miyashita, & W. O'Neil (Eds.), *Image, language, brain* (pp. 149–162). Cambridge, MA: MIT Press.

Ito, M. (2002). Historical review of the significance of the cerebellum and the role of purkinje cells in motor learning. *Annals of the New York Academy of Sciences, 978,* 273–288.

Ito, M. (2008). Control of mental activities by internal models in the cerebellum. *Nature Reviews Neuroscience, 9,* 304–313.

Ito, M., Sakurai, M., & Tongroach, P. (1982). Climbing fiber induced depression of both mossy fiber responsiveness and glutamate sensitivity of cerebellar Purkinje cells. *Journal of Physiology (Lond.), 324,* 113–134.

Jueptner, M., & Weiller, C. (1998). A review of differences between basal ganglia and cerebellar control of movements as revealed by functional imaging studies. *Brain, 121,* 1437–1449.

Justus, T. (2004). The cerebellum and English grammatical morphology: Evidence from production, comprehension, and grammaticality judgments. *Journal of cognitive neuroscience, 16*(7), 1115–1130.

Kalman, R. E. (1960). "A New Approach to Linear Filtering and Prediction Problems" *Transactions of the ASME – Journal of Basic Engineering, 82,* 35–45.

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

Kawato, M., Furukawa, K., & Suzuki, R., (1987). A hierarchical neural-network model for control and learning of voluntary movement. *Biological Cybernetics, 57,* 169–185.

Kempson, R. & Cann, R. (2007). Dynamic syntax and dialogue modelling: preliminaries for a dialogue-driven account of syntactic change. *Current Issues in Linguistic Theory, 284*, 73–101.

Lehmann, C. (1995). *Thoughts on grammaticalization.* Munich: Lincom Europa. (First published as *akup* 48, Institut für Sprachwissenschaft, Universität zu Köln, 1982.)

Lehmann, C. (2004). Theory and method in grammaticalization. *Zeitschrift fur Germanistische Linguistik 32*(2), 152–187.

Leiner, H. C., Leiner, A. L., & Dow, R. S. (1986). Does the cerebellum contribute to mental skills? *Behavioral Neuroscience, 100*(4), 443–54.

Leiner, H. C., Leiner, A. L., Dow, R. S. (1987). Cerebro-cerebellar learning loops in apes and humans. *Italian Journal of Neurological Science. 8,* 425–436.

Leiner, H. C., Leiner, A. L., & Dow, R. S. (1991). The human cerebro-cerebellar system: Its computing, cognitive, and language skills. *Behavioural Brain Research*, *24*, 113–128.

Linebarger, M. C., Schwartz, M. F., & Saffran, E. M. (1983). Sensitivity to grammatical structure in so-called agrammatic aphasics. *Cognition, 13,* 361–392.

MacLeod, C.E., Zilles, K., Schleicher, A., Rilling, J.K., & Gibson, K.R. (2003). Expansion of the neocerebellum in Hominoidea. *Journal of human evolution, 44,* 401–429.

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.

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

Middleton, F. A., & Strick, P. L. (1998). Cerebellar output: Motor and cognitive channels. *Trends in Cognitive Sciences, 2*(9), 348–354.

Mishkin, M., Malamut, B., & Bachevalier, J. (1984). Memories and habits: Two neural systems. In G. Lynch, J. L. McGaugh, & N. M. Weinberger (Eds.). *Neurobiology of learning and memory* (pp. 65–77). New York: Guildford Press.

Molinari, M., Leggio, M. G., & Silveri, M. (1997). Verbal fluency and agrammatism. In J.D. Schmahmann (Ed.), *The cerebellum and cognition* (pp. 325–339). San Diego: Academic Press.

Mugnaini, E. (1983). The length of cerebellar parallel fibers in chicken and rhesus monkey. *Journal of Computational Neurology, 220,* 7–15.

Nelson, M. E., & Paulin, M. G. (1995). Neural simulations of adaptive reafference suppression in the elasmobranch electrosensory system. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology, 177*(6), 723–736.

Nicolson, R. I., Fawcett, A. J., & Dean, P. (2001). Developmental dyslexia: The cerebellar deficit hypothesis. *Trends in Neurosciences, 24,* 508–511.

Paulin, M. G. (1989). A Kalman filter theory of the cerebellum. In M. A. Arbib & S.-I. Amari (Eds.), *Dynamic Interactions in Neural Networks: Models and Data* (pp. 239–259). New York: Springer.

Paulin, M. (1997). Neural representations of moving systems. In J. D. Schmahmann (Ed.), *The Cerebellum and Cognition, International Review of Neurobiology, 41,* (pp. 515–533). San Diego: Academic Press.

Petrides, M., & Pandya, D. (1994). Comparative architectonic analysis of the human and the macaque frontal cortex. In F. Boller & J. Grafman (Eds.), *Handbook of neuropsychology* (Vol. 9, pp. 17–58). Amsterdam: Elsevier.

Pickering, M. J., & Garrod, S. (2007). Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences, 11*(3), 105–110.

Pickett, E. R. (1998). *Language and the Cerebellum.* Doctoral dissertation, Brown University.

Saffran, E. M., Schwartz, M.F., & Linebarger, M. C. (1998). Semantic influences on thematic role assignment: Evidence from normals and aphasics. *Brain and Language, 62,* 255–297.

Samuel, A. G. (1981). Phonemic restoration: Insights from a new methodology. *Journal of Experimental Psychology: General, 110,* 474–494.

Sanford, A. J., and Sturt, P. (2002) Depth of processing in language comprehension: not noticing the evidence. *Trends in Cognitive Sciences, 6*(9), 382–386.

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

Silveri, M. C., Leggio, M. G., & Molinari, M. (1994). The cerebellum contributes to linguistic production: A case of agrammatic speech following a right cerebellar lesion. *Neurology, 44,* 2047–2050.

Sperling, A. J., Lu, Z.-L., Manis, F. R., & Seidenberg, M.S. (2005). Deficits in perceptual noise exclusion in developmental dyslexia. *Nature Neuroscience, 8*(7), 862–863.

Stowe, L.A., Paans, A.M.J., Wijers, A.A., & Zwarts, F. (2004). Activations of "motor" and other non-language structures during sentence comprehension. *Brain and Language, 89,* 290–299.

Thach, W.T. (1998). What is the role of the cerebellum in motor learning and cognition? *Trends in Cognitive Sciences, 2*(9), 331–337.

Thach, W. T., Goodkin, H. P., & Keating, J. G. (1992), The cerebellum and the adaptive coordination of movement. *Annual Review of Neuroscience, 15,* 403–442.

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

Ullman, M.T. (2004). Contributions of memory circuits to language: the declarative/procedural model. *Cognition, 92,* 231–270.

Warren, R.M. (1970). Perceptual restoration of missing speech sounds. *Science, 167*, 392–393.

Wason, P. & Reich, S. S. (1979) A verbal illusion. *Quarterly Journal of Experimental Psychology, 31,* 591–597.

Watkins, K.E., Strafella, A.P., & Paus, T. (2003). Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia, 41,* 989–994.

Whiting, B.A., & Barton, R.A. (2003). The evolution of the cortico-cerebellar complex in primates: anatomical connections predict patterns of correlated evolution. *Journal of Human Evolution, 44,* 3–10.

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

Wolpert, D.M., Miall, R.C., & Kawato, M. (1998). Internal models in the cerebellum. *Trends in Cognitive Sciences, 2*(9), 338–347.

Argyropoulos, G. P. (2010a). Is grammaticalization glossogenetic? In A. D. M. Smith, M. Schouwstra, B. de Boer, & K. Smith (Eds.), *The Evolution of Language: Proceedings of the 8th International Conference on the Evolution of Language* (pp. 3-10). Singapore: World Scientific.

IS GRAMMATICALIZATION GLOSSOGENETIC?

GIORGOS P. ARGYROPOULOS
giorgos@ling.ed.ac.uk
Language Evolution and Computation Research Unit,
University of Edinburgh, EH8 9LL, United Kingdom

It has recently been suggested that grammaticalization can be fruitfully explained by the glossogenetic mechanisms for language evolution and historical change. Contrary to this position, it is here argued that the incorporation of grammaticalization processes in the glossogenetic ontology is far from unproblematic.

1. Introduction: The Glossogenetic Framework

In glossogenetic models of cultural transmission, language changes through misconvergences between the hypothesis generated by Agent 2 and the one based on which Agent 1 outputs the data that Agent 2 eventually observes (e.g., Brighton *et al.*, 2005). The role of learning/processing biases is here catalytic: linguistic input challenging such constraints does not propagate through the bottleneck, and is marginalized/ regularized by more learnable/ processible schemata, instances of which gradually dominate in *'the arena of use'* (Hurford, 1987). In this respect, *'it is languages, not language users that are adapting'*, since the emergence of more functional forms of communication is *'merely a happy byproduct of the adaptive mechanism at work'* (Kirby *et al.*, 2004), or, as an instance of language adapting to the human brain, and not the reverse (Christiansen & Chater, 2008). Glossogenesis thus provides a radical alternative to (strong versions of) generativist evolutionary considerations, according to which language evolution is principally explicable with respect to the biological evolution of a Language Acquisition Device.

2. Grammaticalization beyond Glossogenesis

Among others, Christiansen and Chater (2008) have recently suggested that the same mechanism can explain grammaticalization phenomena. Contrary to this, I would like to argue that the basic glossogenetic premise of language adaptively changing according to learnability and processibility constraints of the human brain is problematic when employed for the explanation of grammaticalization. In particular, it is susceptible to the same arguments addressed against generativist accounts of grammaticalization (see below for references).

2.1. The latent generativist premise

Such accounts have been of paradigmatic value for equivalent evolutionary modeling attempts: Brighton *et al.*, (2005), suggestively, reference simulations of changes resulting *"directly from missconvergences arising during language acquisition"*, such as Clark&Roberts (1993), Niyogi&Berwick (1997), Briscoe (2002), studies explicitly assuming the Principles and Parameters (Chomsky, 1981) framework. Here, the language processor meets significant costs in assigning a particular structure to certain instances of ambiguous input, and reanalyzes them according to a considerably more acquirable/ processible structure, as in Robert's (1993: 228-9) *"Least Effort Strategy"* in the changes of the future tense in Romance (see also Lightfoot (1979) on English modals).

Characteristically, Hashimoto and Nakatsuka (2006) assume Campbell's (2001) heavily criticized position that reanalysis (for Langacker(1977, p. 59), a *'change in the structure of an expression or class of expressions that does not involve any immediate or intrinsic modification of its surface manifestation'*) and analogy (according to Hopper and Traugott (1993, p. 56), *'the attraction of extant forms to already existing constructions'*) underlie grammaticalization, and instantiate them in the operations that linguistic agents perform in learning and generalizing their grammar in Kirby's (2002) compositionality model. However, reanalysis has been argued to be dissociable from grammaticalization: reanalysis is an abrupt phenomenon, does not involve the loss of autonomy of linguistic signs, is not inherently unidirectional, and presupposes ambiguity and mis-processing/ mislearning of the input; Reanalysis occurs without the necessary accompaniment of grammaticalization, and, conversely, grammaticalization may occur without

reanalysis. Reanalysis involves contracting new syntactic relations with sentence elements not related before, whereas grammaticalization does not involve any real change in constituent structure. Grammaticalization may lead to structures and categories that have not existed, while reanalysis, acting on an analogical basis, only operates on already available categories (see Haspelmath(1998) and Lehmann (2004) for a discussion of the above points) and thus may not suffice as an explanation for the development of a new grammatical category. In this sense, input ambiguity and the interaction of misconvergences with capacity limitations are explicitly considered irrelevant with grammaticalization processes. The argument remains the same if one is to employ *'processing preferences'*-based explanations, (such as Hawkins' (2004), explicitly distinguishes grammaticalization processes from the ones in his work) to account for grammaticalization operations: The derivation of a new grammatical category is not explicable by the adoption of a more preferable template (already there) for the analysis of a given input string.

It has similarly been suggested that semantic miscovergence in historical change might provide a glossogenetic insight into grammaticalization. However, what has been termed *'semantic reanalysis'* (Hoefler and Smith (2009) use the term *'reanalysis'* to denote what Heine (2003) has called *'context-induced reinterpretation'*) is far from intrinsic to grammaticalization, as it pertains to a radically broader range of lexical items, only a minority of which acquires grammatical functions. Despite allowing expressions to increase in frequency of occurrence and acquire basic discourse status, it does not explain the core transition from lexical-conceptual to grammatical-procedural status that grammaticalizing expressions undergo (Nicolle, 1998; Bybee et al., 1994; Haspelmath, 1999).

Thus, despite its emphasis on E-language properties, glossogenesis crucially remains, like the generativist explanation of grammaticalization, a *"competence-based"* (Haspelmath, 1998) one: it is committed to viewing E-language properties as explanatorily significant only for the negotiation of I-language representations, and not for the introduction of *'performance changes'* (*ibid.*), the results of which macroscopically provide grammaticalized variants.

2.2. ***Ontological dissociation: Reducing and Conserving***

In delimiting the ontology of grammaticalization and glossogenetic mechanisms, it is valuable to consider the distinction between the *'Conserving'* and the *'Reducing effects'* of linguistic repetition in historical language change (Bybee & Thompson, 2000): Glossogenesis is best construable as studying the *'Storage effects'* (*ibid.*) of repetition (conservation of irregularity, analogical leveling-generalization-regularization, reanalysis), not the *'Processing effects'* (*ibid.*). In their brief discussion of grammaticalization processes, suggestively, Christiansen and Chater (2008) reference Hare and Elman's (1995) simulation as an example of historical language change studied on the basis of learning/processing constraints. Yet the phenomenon described is one of analogical leveling and conservation phenomena in the acquisition of verbal morphology, and is far from relevant with grammaticalization processes to encourage a similar approach to the latter. Hawkins' (2004) processing principles, or even Batali's (2002) work on exemplar competition provide characteristic cases (the ones consistent with observations become liable for re-usage and creation of new S-M mappings, while others are used less often and are overtaken by the most prominent ones).

Grammaticalization, on the other hand, precisely represents a multilevel *"reducing effect"* (Bybee & Thompson, 2000) of repetition. Automatization of performance in language processing, the most widely supported cognitive core of grammaticalization (e.g., Givòn, 1979; Bybee, 1998; Haspelmath, 1999; Lehmann, 2004), represents an instance of domain-general, non-species-specific adaptive responses against repeated behavioral repertoires, and involves particular neurocognitive mechanisms (see Argyropoulos (2008) for neurolinguistic reflections). Here, linguistic change does not involve misprocessing of ambiguous input, but is the result of the user's adaptive minimization of cognitive and attentional/motoric costs, and hence of the freedom in the manipulation of linguistic signs (Lehmann, 2004). A clear example of the dissociation of those two kinds of change (summarized in **Figure 1**) is shown in Hooper (1976), where sound change is demonstrated to affect high-frequency items first, and analogical leveling to affect low-frequency items first.

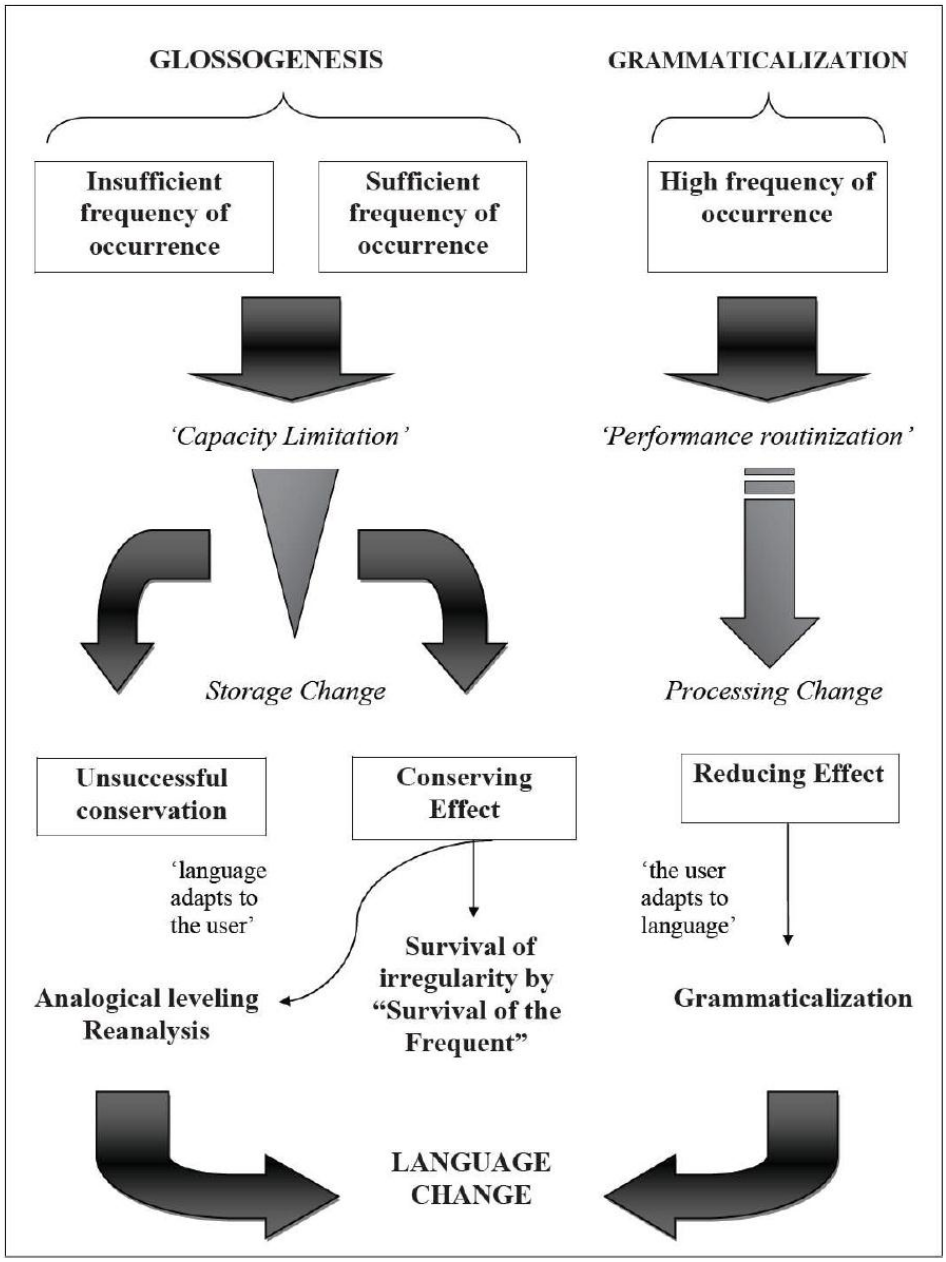


Figure 1: Dissociation between the ontology of glossogenesis and the one of grammaticalization, following Bybee & Thompson's (2000) distinction between the conserving and reducing effects of linguistic repetition; despite the rich interactions between the two effects (not portrayed here), both the changes and their corresponding triggers are clearly dissociable.

3. Functionalist Criticism and Directionality of Language Transmission

Functionalist criticism (Croft, 2004) on glossogenesis has failed to address this topic, rather emphasizing the need to abandon vertical (language acquisition-based) models of transmission for the simulation of historical change; but this premise has been identified as only a methodologically justified simplification (Smith *et al.*, 2003). Yet even in a horizontal transmission model, change is introduced with agents misattributing an underlying representation to linguistic strings because of processing limitations encountered. In Batali (2002, p. 115), for example, where the sender transmits a signal S to express a meaning M1, following a particular analysis A1 of the mapping from the structure of M1 to the signal S, with the receiver deriving an interpretation M2 according to analysis A2 of the mapping from the signal structure to M2, language change is initiated by a discrepancy between A1 and A2, where A2 is more optimal with respect to the processing constraints of the language agent than A1. This is the case in reanalysis and analogical leveling, or in the conservation of irregular frequent patterns- not in grammaticalization, unless one adheres to heavily criticized generativist constraint-based modeling.

4. Concluding remarks

The issue addressed here apparently does not refer to the feasibility of computational simulations of grammaticalization processes, but questions the compatibility of glossogenesis with the explanandum of grammaticalization. Thus, the eventual convergence of grammaticalization theory and evolutionary computational modeling (Hurford, 2003) still deserves encouragement; however, according to the strongest version of the argument, such convergence is impossible in the glossogenetic framework, as its ontology is by definition distinct from grammaticalization phenomena; in its weakest version, the argument instructs that such contradictory assumptions be identified and disposed of in this interdisciplinary discourse.

Acknowledgments

The author warmly acknowledges the financial support from the Pavlos and Elissavet Papagiannopoulou Foundation.

References

Argyropoulos, G. (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 Press.

Batali, J. (2002). The negotiation and acquisition of recursive grammars as a result of competition among exemplars. In E. Briscoe (Ed.) *Linguistic Evolution through Language Acquisition: Formal and Computational Models* (pp. 111–172). Cambridge: Cambridge University Press.

Brighton, H., Kirby, S., & Smith, K. (2005). *Cultural selection for learnability: Three hypotheses underlying the view that language adapts to be learnable.* In Tallerman, M. (Ed.) *Language Origins: Perspective on evolution.* Oxford University Press, Oxford.

Briscoe, E. (Ed.) (2002). *Linguistic Evolution through Language Acquisition: Formal and Computational Models.* Cambridge: Cambridge University Press.

Bybee, J. L. (1998). A functionalist approach to grammar and its evolution. *Evolution of Communication, 2,* 249-278.

Bybee, J.L., Perkins, R., & Pagliuca, W. (1994). *The evolution of grammar. Tense, Aspect, and Modality in the Languages of the World.* Chicago and London.

Bybee, J. L. & Thompson, S. (2000). Three frequency effects in syntax. *Berkeley Linguistic Society, 23,* 65-85.

Campbell, L. (2001). What's wrong with grammaticalization? *Language Sciences, 23*, 113-161.

Chomsky, N. (1981). Principles and parameters in syntactic theory. In N. Hornstein & D. Lightfoot (Eds.), Explanations in Linguistics. London: Longman.

Christiansen, M.H., & Chater, N. (2008). Language as shaped by the brain. *Behavioral and Brain Sciences*, *31(5)*, 489-509.

Clark, R., & Roberts, I. (1993). A computational model of language learnability and language change. *Linguistic Inquiry, 24*, 299-345.

Croft, W. (2004). Form, meaning and speakers in the evolution of language. *Studies in Language*, *28(3)*, 608-611.

Givòn, T. (1979). *On Understanding Grammar*. New York: Academic Press. Structure. Chicago: University of Chicago.

Hare, M., & Elman, J.L. (1995). Learning and morphological change. *Cognition, 56,* 61-98.

Hashimoto, T., & Nakatsuka, M. (2006), Reconsidering Kirby's compositionality model toward modeling grammaticalization. In A. Cangelosi et al. (Eds.), *"The Evolution of Language, Proceedings of the 6th International Conference (EvoLang6),* (pp. 415-416). World Scientific.

Haspelmath, M. (1998). Does grammaticalization need reanalysis? *Studies in Language, 22*, 315-51.
Haspelmath, M. (1999). Why is grammaticalization irreversible? *Linguistics 37(6),* 1043-1068.
Hawkins, J. A. (2004). *Efficiency and complexity in Grammars*. Oxford: Oxford University Press.
Heine, B. (2003). 'Grammaticalization.' In B. D. Joseph, & R. D. Janda (Eds.) *The Handbook of Historical Linguistics*. (pp. 575-601) Oxford: Blackwell.
Hoefler, S., & Smith, A.D.M. (2009). The Pre-linguistic Basis of Grammaticalisation: A Unified Approach to Metaphor and Reanalysis. *Studies in Language* 33(4), 883-906.
Hooper, J. B. (1976). Word frequency in lexical diffusion and the source of morphophonological change. In W. Christie (Ed.). *Current progress in historical linguistics* (pp. 95-105). Amsterdam: North Holland.
Hopper, P., & Traugott, E.C. (1993). *Grammaticalization*. Cambridge: Cambridge University Press.
Hurford, J. (1987). *Language and Number: The Emergence of a Cognitive System*. Oxford: Basil Blackwell.
Hurford, J. (2003). The Language Mosaic and its Evolution. In M.H. Christiansen and S. Kirby, (Eds.), *Language Evolution: The States of the Art*. Oxford University Press.
Kirby, S. (2002). Learning, bottlenecks, and the evolution of recursive syntax. In E. J. Briscoe (Ed.), *Linguistic Evolution through Language Acquisition*. Cambridge: Cambridge University Press.
Kirby, S., Smith, K., and Brighton, H. (2004). From UG to Universals: Linguistic adaptation through iterated learning. *Studies in Language, 28(3),* pp. 587- 607.
Langacker, R. (1977). Syntactic reanalysis. In C. Li (Ed.) *Mechanisms of Syntactic Change* (pp. 57–139). Austin, TX: University of Texas Press.
Lehmann, C. (2004). Theory and method in grammaticalization. *Zeitschrift fur Germanistische Linguistik, 32(2),* 152-187.
Lightfoot, D. (1979). *Principles of Diachronic Syntax*. New York: Cambridge University Press.
Nicolle, S. (1998), A relevance theory perspective on grammaticalization. *Cognitive Linguistics, 9(1)*, 1-35.
Niyogi, P., & Berwick, R. (1997), Evolutionary consequences of language learning. *Linguistics and Philosophy, 20*, 697-719.
Roberts, I. (1993). A formal account of grammaticalization in the history of Romance futures. *Folia Linguistica Historica ,13(1-2)*, 219-258.
Smith, K., Kirby, S. Brighton, H. (2003). Iterated Learning: A framework for the emergence of language. *Artifical Life, 9(4)*, 371-386.

Argyropoulos, G. P. (2010b). *Theta-burst stimulation of the right neocerebellar vermis selectively enhances lexical associative priming.* Manuscript submitted for publication.

Theta Burst Stimulation of the Neocerebellar Vermis selectively enhances Lexical Associative Priming

Giorgos P. Argyropoulos
Language Evolution and Computation Research Unit, Linguistics and English Language, School of Philosophy, Psychology and Language Sciences, University of Edinburgh, Dugald Stewart Building, 3 Charles Street, Edinburgh EH8 9AD, United Kingdom
(E-mail: giorgos@ling.ed.ac.uk)

Argyropoulos GP. Theta burst stimulation of the neocerebellar vermis selectively enhances lexical associative priming. *J Neurophysiol* …Recent research in cerebellar cognitive and linguistic functions makes plausible the idea that the cerebellum is involved in processing temporally contiguous linguistic input. In order to assess this hypothesis, a simple lexical decision task was constructed to study whether the effect of transcranial magnetic stimulation on two different neocerebellar sites would have a selective impact on associative as opposed to semantic priming. This is the first experiment applying transcranial magnetic stimulation of the cerebellum to a linguistic task. The results show a selective increase of associative priming sizes after stimulation of the neocerebellar vermis that cannot be attributed to changes in sensorimotor performance or in accuracy rates. The finding is discussed within the context of associative cerebellar computations and neocerebellar vermal pathology in schizophrenia.

INTRODUCTION

Language perception and comprehension has been recently proposed to dynamically combine the outputs of an 'input analysis system' with predictions of internal forward models, generating an estimate of the next state that the input analysis system will enter (Pickering & Garrod, 2007). The cerebellum (CB) is well established to instantiate computations of state estimation (Paulin, 1989), internal models (Wolpert et al., 1998), and, generally, predictive mechanisms for temporally contiguous events occurring in the central nervous system (Courchesne & Allen, 1997). The cytoarchitectural homogeneity of the CB has supported the idea of a uniform, domain-general CB computation, including language processing (e.g., Schmahmann, 2000). This idea is summarized in the Neocerebellar Kalman Filter Linguistic Processor (NCBKFLP) hypothesis (Argyropoulos, 2009), whereby computations of state estimation in language perception-comprehension are fundamentally undertaken by neocerebellar (NCB) circuitry. At an elementary level, the hypothesis would predict that priming phenomena in processing 'phrasal associates' (Hutchison, 2003), e.g., 'battery-HEN', 'skeletons-CLOSET', should involve NCB predictive computations. This would be because associative relatedness reflects word use, i.e., the probability that one word will call to mind a second one (e.g., Thompson-Schill et al., 1998). On the other hand, priming phenomena not relying on the temporal contiguity of linguistic input should not directly involve CB circuitry. This would be the case for what is traditionally termed 'semantic priming': semantic relatedness reflects featural overlaps of the semantics of the two words, e.g., 'car-TRUCK',

'swan-CHICKEN' (*ibid*). Research in Alzheimer's disease and semantic dementia has established that temporal lobe lesions induce disruptions in semantic, but not in associative priming (e.g., Rogers & Friedman, 2008). While the reverse dissociation remains elusive, indirect evidence supports NCB involvement in associative linguistic processes. Associative learning, from eyeblink conditioning (Woodruff-Pack et al., 1996) to matching colours to numerals (Drepper et al., 1999) is deficient in CB patients. Right posterolateral cerebellar (PLCB) patients have shown difficulties in generating appropriate verbs for noun-stimuli (e.g.., 'chef-COOKING'); however, their inappropriate responses are still semantically related with the noun-stimuli (e.g., 'pill-SMALL'), indicating that CB lesions do not compromise semantic networks per se (Fiez et al., 1997).

Lexical Decision Task

Stimuli

Each session contained 600 trials (word/nonword-word/nonword) pairs, and was divided in two blocks of 300 trials each. Each block consisted of two 'mini-blocks' of 150 trials. Mini-blocks contained 60 word-word, 60 nonword-nonword, 15 nonword-word, and 15 word-nonword trials. This yielded an equal number of 'yes' and 'no' responses, as well as an equal number of words and nonwords. In the second and fourth mini-blocks, the same word and nonword pairs were presented as those of the first and third mini-blocks, respectively. This enabled to study changes in practice-related effects after CB TMS. In order to avoid onset effects (Forster & Davis, 1991), primes and targets always differed in the first phoneme; they were also not orthographically similar, and the graphemic structure of nonwords resembled that of orthographically legal words. Generation of categorically subordinate terms for a generic term (e.g., 'fruit-APPLE') is similarly not compromised in PLCB pathology, suggesting that the CB is not involved in processing semantic-categorical relations (Gebhart et al., 2002). The present experiment thus set out to explore the effects that transcranial magnetic stimulation (TMS) on the NCB would have on associative and semantic priming in a lexical decision task. Repeated trials were also introduced, in order to study the well-documented CB involvement in practice-induced facilitation in processing repeated stimuli (e.g., Fiez et al., 1997; Gebhart et al., 2002; Ferrucci et al., 2008).

METHODS

Printed frequency was taken from the British National Corpus (BNC, written part). Prime and target words did not differ in frequency or length across blocks or mini-blocks [mean prime frequency: 32.82 words/million; SD: 41.72; mean target frequency: 38.21; SD: 41.68; mean prime length: 5.73 letters; SD: 1.58; mean target length: 5.58; SD: 1.71; all *Fs*, $F<1$]. The first and third mini-blocks contained 15 related pairs and 15 unrelated pairs for each of the two different priming sets: semantically related pairs, where the prime word is a subordinate term of that in the target position, but does not form a phrase in speech with it (e.g., 'apple-FRUIT'); and phrasal associates, where the prime word co-occurs in speech with but is categorically unrelated to the target word (e.g., 'battery-HEN'). Two lists were created, across which subjects were rotated. The stimuli of the second half of the first list were the stimuli of the first half of the second list, and vice versa (table 1). The relatedness proportion, i.e., the ratio of related per unrelated words, was 0.2, as in

other experiments (e.g., Rogers & Friedman, 2008). Such low proportion minimizes predictability and maintains an automatic level of priming, given that strategic components may be introduced otherwise (Ober, 2002).

As in other lexical decision tasks, target words of the associative group were not matched with those of the semantic group (e.g., Ferrand & New, 2003), in that the former were shorter [STIMULUS SET: $F(1,112) = 11.62$; $P < 0.005$]. Word frequency and length for primes and targets were thus matched across mini-blocks separately for each priming type [Associative set: related pairs: mean prime frequency: 34.73 words/million; SD: 45.36; mean prime length: 5.72 letters; SD: 1.64; mean target frequency 37.45 words/million; SD: 34.21; mean target length: 5.2 letters; SD: 1.44; unrelated pairs: mean prime frequency: 32.8 words/million; SD: 43.82; mean prime length: 5.55 letters; SD 1.49; mean target frequency: 37.45 words/million; SD: 34.21; mean target length: 5.2 letters; SD: 1.44. Semantic set: related pairs: mean prime frequency: 31.10 words/million; SD: 30.97; mean prime length: 5.85 letters; SD: 1.70; mean target frequency: 38.96 words/million; SD: 48.64; mean target length: 5.95 letters; SD: 1.86; unrelated pairs: mean prime frequency: 32.63 words/million; SD: 44.74; mean prime length: 5.8 letters; SD: 1.54; mean target frequency: 38.96 words/million; SD: 48.64; mean target length: 5.97 letters; SD: 1.90; all Fs, $F<1$]. Associative strength was determined on the basis of free word-generation norms provided in the Edinburgh word Association Thesaurus (EAT; Kiss et al., 1973), and by the frequency of co-occurrence of those words in the BNC. All associated pairs co-occurred in speech in an immediate fashion. For example, a pair like 'battery-HEN' was selected, but not a pair like 'storm-TEACUP'. Associated pairs co-occurred as phrases much more frequently in the BNC than the semantic ones across the two mini-blocks [Associative set: mean strength: 0.56 phrases/million; SD: 0.80. Semantic set: mean strength: 0.01 phrases/million; SD: 0.02; STIMULUS SET: $F(1,56) = 15.23$; $P < 0.000$; rest of Fs, $F<1$], and had higher word association values [Associative set: mean: 0.17; SD: 0.06. Semantic set: mean: 0.01; SD: 0.02; STIMULUS SET: $F(1,56) = 165.83$; $P < 0.000$; rest of Fs, $F<1$]. Semantically related pairs were drawn from the published category norms by McEvoy and Nelson (1982), and Van Overschelde et al. (2004); they were further assessed on the grounds of their semantic similarity, which was determined using the 'WordNnet: Similarity' software (Pedersen et al., 2004). Semantically related pairs received much higher similarity ratings than the associatively related ones across the two mini-blocks [Associative set: mean: 0.12; SD: 0.03. Semantic set: mean: 0.42; SD: 0.09; STIMULUS SET: $F(1,56) = 263.12$; $P < 0.000$; rest of Fs, $F<1$].

Table 1. Task conditions with example stimuli.

Procedure

Prime and target stimuli were presented in green colours at the centre of a black screen of a MacBook using the 'DirectRT' (Jarvis, 2008) software. Each trial consisted of the following sequence of three stimuli presented on the same screen location. First a fixation point (a cross '+') was presented for 400 ms, followed by presentation of the prime in lowercase letters for 100 ms, which was followed by the presentation of the target in uppercase letters. The targets remained on the screen until participants responded. As no masking

mediated the presentation of the prime and the target, the SOA value was confounded with that of the prime duration, i.e., 100 ms. SOAs longer than 200-250 ms are generally considered to involve strategic effects (e.g., Perea & Rosa, 2002), and were thus not preferred. There was an intertrial interval of 600 ms. Participants were instructed to focus on the fixation point, read the prime, and respond only to the target. The order of stimulus presentation was randomized for each subject. Subjects were instructed to press one of two buttons on the keyboard ('j' for yes and 'f' for no) to indicate whether the target letter string was an English word or not, as rapidly and accurately as possible. They used their dominant right hand index finger for the word responses, and their left hand index finger for the nonword responses. Each subject received a total of 20 practice trials prior to the 600 test trials. The whole session lasted approximately 45 minutes. Participants were tested individually in a silent and dimly lit room. They received written instructions explaining the task; they also provided informed consent and received compensation for each session.

Pilot experiment

In order to establish that related trials yielded the necessary priming effects, 18 subjects were recruited to participate in one session, with no intervention of TMS apparatus. Their latencies were subjected to a three-way (PRIMING TYPE, PHASE, RELATEDNESS) within-subjects ANOVA (figure 1, table 2). Related trials received significantly shorter latencies than unrelated ones across the two priming types [RELATEDNESS: $F(1,17) = 12.07$; $P < 0.005$]. Associative priming was marginally stronger [PRIMING TYPE X RELATEDNESS $F(1,17) = 2.80$; $P = 0.11$]. This was expected, given the overall trend in lexical decision experiments for semantic priming to be weaker than associative priming (Lucas, 2000). These effects occurred independently of the phase of the session [$F<1$].

FIG. 1. Mean latencies per unrelated and related items for pilot group. Error bars represent + 1 MSe.

Table 2. Mean RTs per condition for pilot group

TMS

TMS protocol

TMS was delivered via a 70-mm figure-eight coil connected to a Magstim Super Rapid Transcranial Magnetic Stimulator (Magstim Company, Whitland, UK). The rTMS protocol used was continuous theta burst stimulation (cTBS; Huang et al., 2005), employing a brief burst of 3 low-intensity, high-frequency (50 Hz) TMS pulses every 200 ms, delivered at a 5 Hz rhythm. The experiment employed the offline cTBS procedure: after completion of the first half of each session, 40 seconds (600 pulses) of cTBS were applied on the CB. Stimulus intensity was fixed across participants at 45% of maximum machine output. The coil was positioned tangentially to the scalp, with the handle pointing superiorly. The current in the coil was directed upward, which induced downward current in the CB cortex. This coil position has been found to be optimal for suppressing the contralateral motor cortex in single-pulse TMS (e.g., Oliveri et al., 2005) and to interfere with cognitive processes such as procedural learning in 1 Hz rTMS paradigms (e.g., Torriero et al., 2007).

TMS localization

A medial and a lateral CB site were stimulated. A disadvantage was that no neuronavigational software was used to individualize the exact site of stimulation for each participant; the sites were instead defined on the basis of scalp-based coordinates, as in many other CB TMS studies (e.g., Miall & Christensen, 2004; Oliveri et al., 2005; Koch et al., 2008). The medial site was the right neocerebellar vermis (NCBV), which comprises the right posterior superior compartments of the vermis, and was located 1 cm below and 1 cm laterally to the right from the inion. Stimulation of the NCBV with these coordinates has induced behavioral effects with high spatial precision (Hashimoto & Ohtsuka, 1995). The other site was located 1 cm below and 4.5 cm laterally to the right from the inion, targeting a posterolateral cerebellar site (PLCB). In order to estimate the depth of each site and thus the possibility of their successful stimulation, a volunteer was employed whose brain image was already registered with the Brainsight TMS-MRI co-registration system (Rogue Research, Montreal, Canada). The two scalp-based coordinates used here were transformed into Talairach brain coordinates. For the NCBV site, transformation yielded the coordinates of x = 9 mm, y = -87 mm, z= -21 mm. As shown in figure 2, the medial site is thus successfully localized below the primary fissure and above the prepyramidal fissure, corresponding to the decline, folium, and tuber in the right superior posterior CB vermal lobules VI and VII (Larsell & Jansen, 1972). For the PLCB site, located between the same fissures, the transformation yielded the coordinates of x = 33 mm, y = -87 mm, z = -27 mm, corresponding to the significantly deeper right posterior superior hemispheric lobule HVI (fig. 3).

FIG. 2. Localization of the right NCBV site.

FIG. 3. Localization of the right PLCB site.

Subjects

All subjects (n =8) were right-handed native speakers of English, with normal or corrected-to-normal vision, and with no known attention/ motor/reading deficits (mean age: 26.9 years; min: 19; max: 43; SD: 8.6). They underwent TMS in alternate sites (PLCB, NCBV) in two different sessions on two different days (mean time distance: 13.5 days; min 3; max 26; SD: 9.1), performing the exact same task. Ethical approval was granted by the local ethics committee. All subjects provided their informed consent, and all tolerated CB TBS well, reporting a mild discomfort due to muscle twitching.

Data analysis and statistics

Subjects were rotated across the two different stimulus lists, and across the two possible orders in which the two sites could be stimulated. A 2x2x2 within-subjects design was thus employed, with the following variables: PHASE (before/after TBS), SITE (NCBV/ PLCB), PRIMING TYPE (associative/semantic). The dependent variable was the size of priming for both priming types. Following the standard convention, the measure of the priming effect was the RT mean for the unprimed target words minus that for the primed target words (Ober, 2002). RTs were log-transformed, and any trials with latencies longer than 1200 ms (0.4% of the data) were omitted from the analysis, as reflecting the subject's low familiarity with the stimulus or distraction, rather than lexical access (Perea & Gotor, 1997).

Accuracy rates were arcsine transformed to stabilize variances (e.g., Moss et al., 1995).

RESULTS

TMS on the NCBV induced a selective enhancement in associative priming (table 3, figure 4). A three-way ANOVA shows a significant PRIMING TYPE X SITE X PHASE interaction [$F(1,7) = 20.8$; $P < 0.005$]. For the conditions of NCBV stimulation, a two-way ANOVA shows a significant PRIMING TYPE X PHASE interaction [$F(1,7) = 10.08$; $P < 0.05$], with no change in priming sizes independently of a priming type [PHASE: $F<1$]. No significant change is found after PLCB TMS, selectively for a priming type or not [PRIMING TYPE X PHASE: $F(1,7) = 2.85$; $P = 0.14$; PHASE: $F<1$]. Priming sizes do not differ in the phase before TMS between the two sites [$F<1$], with an only marginal PRIMING TYPE X SITE interaction [$F(1,7) = 3.07$; $P = 0.12$]. On the contrary, a two-way ANOVA for the priming sizes after TMS shows a strong PRIMING TYPE X SITE interaction [$F(1,7) = 15.86$; $P < 0.01$], with an only marginal difference between the two sites in overall priming sizes [SITE: $F(1,7)= 3.32$; $P= 0.11$]. A two-way ANOVA for associative priming sizes shows a strong SITE X PHASE interaction [$F(1,7) = 12.57$; $P < 0.01$], with no differences across sites or phases overall [SITE: $F(1,7) = 1.90$; $P = 0.2$; PHASE: $F<1$]. This interaction is owed to the increased size of associative priming in the condition after NCBV TMS: a one-way ANOVA comparing the two sites of stimulation with respect to associative priming after TMS shows larger associative priming size for NCBV conditions [SITE: $F(1,7) = 14.71$; $P < 0.01$]. Similarly, the only significant priming difference noted between the phase before and after TMS is that for associative priming with NCBV TMS [$F(1,7) = 6.4$; $P < 0.05$]. No change is introduced by PLCB TMS [PHASE: $F(1,7) = 1.80$; $P= 0.22$], and no difference is observed in associative priming sizes before TMS between the two sites [SITE: $F<1$]. A two-way ANOVA for semantic priming also shows a significant SITE X PHASE interaction [$F(1,7) = 9.58$; $P < 0.05$; rest of Fs, $F < 1$]. However, the interaction relies on only marginal differences in semantic priming sizes: the condition before NCBV TMS shows marginally larger semantic priming than that before PLCB TMS [SITE: $F(1,7) = 2.75$; $P = 0.14$]; semantic priming after PLCB TMS is also marginally larger than that after NCBV TMS [SITE: $F(1,7) = 3.42$; $P = 0.11$]; there is also a marginal reduction of semantic priming size after NCBV TMS [PHASE: $F(1,7)= 2.16$; $P = 0.19$]; no change introduced in semantic priming after PLCB TMS [PHASE: $F<1$]. Additional analyses were performed, using ratios instead of differences, and/or excluding the 1200 cut-off point; they all showed equivalent effect sizes and significance.

FIG. 4. Priming sizes before and after TMS on the two sites for the two primng types. Error bars represent + 1 MSe.

Table 3. Priming sizes before and after TMS on the two sites for the two priming types.

In an equivalent analysis of accuracy rates, no effects reached significance [PRIMING TYPE X SITE X PHASE: $F < 1$]. Likewise, a three-way analysis of overall lexical decision latencies (WORDNESS, SITE, PHASE) only showed slower latencies for nonwords across sites and phases [WORDNESS: $F(1,7) = 8.33$; $P < 0.05$]

and an expected acceleration of decisions in the second phase across sites and target types [PHASE: $F(1,7)= 8.90$; $P<0.05$; rest of Fs, $F<1$]. An analysis of overall lexical decision accuracy rates showed an only marginal WORDNESS X SITE X PHASE interaction [$F= 3.30$; $P= 0.12$], and no SITE X PHASE interaction [$F<1$]. However, this marginal interaction is not owed to any particular phase before or after stimulation [both Fs, $F<1$]. In order to capture any effects of TMS in processing repeated trials, the mean latencies and accuracy rates for the first and third mini-blocks were subtracted from those for the second and fourth mini-blocks, respectively. A three-way analysis (WORDNESS, SITE, PHASE) on latencies showed only a difference between the two sites across phases [SITE: $F(1,7) = 3.64$; $P = 0.1$; rest of Fs, $F<1$].An analysis for accuracy rates showed a marginal SITE X PHASE interaction [$F(1,7) = 4.31$; $P = 0.08$], that relies though on differences in baseline conditions [SITE: $F(1,7) = 3.9$; $P = 0.09$], but not on any differences between conditions after TMS [SITE: $F<1$].

DISCUSSION

The results above demonstrate that cTBS over the NCBV selectively enhanced associative priming. This is hardly reducible to any effects of TMS on sensory (reading), motor (button key pressing), or elementary cognitive processes (e.g., association between lexical decision, i.e., 'yes/no', and button key-press, i.e., 'j/f'): all such explanations would predict that both related and unrelated pairs of both priming types would involve slower or faster lexical decisions; furthermore, no such effects were observed in the first place. The absence of effect on priming sizes after cTBS of the PLCB site is interpretable in different ways. While the scalp coordinates for the NCBV are precise enough, the CB hemispheres constitute a much broader area. Moreover, the PLCB site was significantly deeper than the NCBV one; indeed, the vermis is closest to the TMS stimulation coil (e.g., Miall & Christensen, 2004), while PLCB stimulation with a cone coil is often preferred (e.g., Miall & King, 2008). Thus, the possibility that the PLCB is equally involved in associative relations can not be excluded. Similarly, no changes were found after TMS in processing repeated trials. This would suggest that the CB sites targeted are irrelevant with practice-related mechanisms, unlike other right CB hemispheric sites, as those found in clinical (Fiez et al., 1997; Gebhart et al., 2002) or in stimulation studies (Ferrucci et al., 2008).

At the neurofunctional level, the enhancement in associative priming would need to be explained in terms of disruptions in Purkinje cell activity; this is because CB stimulation with flat coils most probably stimulates the first layers of the CB cortex (e.g., Oliveri et al., 2005). The inhibitory outputs of Purkinje cells to the deep nuclear neurons serve to modulate the level of the excitation the latter undergo by input from mossy and climbing fiber collaterals. Depending on the prevalence of either nuclear or cortical activity, the CB input to the cerebral cortex may modulate tonically, enhancing excitability, or phasically, reducing the excitability of cerebral cortical outputs (Molinari et al., 2002). Thus, TMS here may have introduced neural noise to the inhibitory activity of the Purkinje cell layer, allowing inappropriate levels of excitatory output to the cerebral cortical targets of the deep CB nuclei. Indeed, 1Hz CB rTMS over the left CB cortex has been found to increase intracortical facilitation of the right motor

cortex, with such low-frequency rTMS decreasing Purkinje cell inhibition on deep CB nuclei, and thus ultimately enhancing thalamocortical facilitation of cerebral cortical outputs (Oliveri et al., 2005). However, while cTBS in principle shows the same depressing effects on cortical activity with 1 Hz rTMS protocols (e.g., Huang et al., 2005), the first results of cTBS on the CB have shown effects in the very opposite direction: MEP amplitude obtained from the contralateral primary motor cortex decreased after cTBS on the CB (Koch et al., 2008). Thus, it remains an open question whether 1 Hz rTMS stimulation would induce similar enhancements or disruptions of associative priming. Therefore, the explanation of the particular direction of the TMS effect remains unclear; rather, the selective nature of the enhancement for both stimulation site and priming type implicates the NCBV in associative linguistic mechanisms.

This finding stands at the intersection of two largely independent strands of research in schizophrenic cognitive deficits, one on selective NCBV, and not PLCB volume reduction (e.g, Okugawa et al., 2003, 2007), and the other on superfluous priming effects in such populations (e.g., Picard et al., 2008). Impaired vermal Purkinje cell inhibitory output to the fastigial nuclei has been considered to induce hyperactivity in dopaminergic basal ganglionic outputs, which may in turn deregulate frontal cortical function via the corticostriatal loop (Martin & Albers, 1995). Similarly, given the connection of the vermis via the fastigial nucleus with the thalamus, the limbic system, and the frontal lobes, vermal abnormality is thought to disrupt those interconnected areas and thus indirectly cause cognitive function to be out of tune (Ichimiya et al., 2001).

However, there is reason to believe that the effects induced here reflect a direct disruption of NCB-based language functions, rather than an indirect cognitive deregulation. The latter would suffice as an explanation only if TMS on the NCBV had induced an enhancement in priming for both priming types. Furthermore, non-motor-related NCBV activations have been reported in a stem-completion task (Desmond et al., 1998), selectively for the condition where the stem involved a small number of appropriate completions (e.g., 'PSA_'). In the same vein, prime words in the associatively related pairs here co-occur in language with a large number of lexical items, but form phrasal idiomatic units only with a very restricted subset. Both findings then would point towards the function of the CB as a 'look-up table' for mappings developed and applied by the significantly slower cortico-cortical processes (Doya, 1999). In this respect, the very computations underlying associative priming are identical with the associative computations of the CB in toto; on a schematic level, they involve the generation of predictions for experiential input 'B' based on its reliable co-occurrence with input 'A' (see introduction). Furthermore, while boosted semantic priming has been observed in schizophrenia (see Picard et al., 2008, for a meta-analysis), there is also evidence suggestive of selective abnormalities in associative priming. In a recent study, healthy controls were primed more for lexical decisions on target words related with primes via two relationships ('semantic-and-associated', e.g., 'doctor-NURSE') than for pairs related only semantically (e.g., 'deer-PONY') or only associatively (e.g., 'bee-HONEY'); on the contrary, schizophrenics showed greater priming for associated-only pairs than for those related only semantically or both semantically and associatively (Nestor et al., 2006). This pattern was attributed to

'an associational bias that leads to restricted semantic integration and contributes to disturbed thinking' (ibid, p. 142). Such NCBV associative deficits could thus underlie the failure of schizophrenics to grasp the essence of a statement and their tendency to respond to more superficial aspects of a conversation (Moritz et al., 2001).

The present study thus provides the first linguistic TMS experiment on the CB, along with evidence for involvement of NCB circuitry in associative linguistic mechanisms. Future research with different TMS protocols would be required to clarify the neurofunctional details of the effect observed, and more powerful coils should be used to investigate the involvement of deeper lateral CB loci.

Acknowledgments

G.P.A would like to warmly thank Dr. Neil Muggleton for subject recruitment and technical support; Dr. Gorana Pobric, for earlier pilot tests; Dr. Patrick Sturt and Dr. Thomas Bak, for helpful comments, and the 'Pavlos and Elissavet Papagiannopoulou Foundation', for the financial support.

REFERENCES

Argyropoulos GP. Neocerebellar emulation in language processing. In: *Brain Talk: Discourse with and in the brain. Papers from the first Birgit Rausing Language. Program Conference in Linguistics,* edited by Alter K, Horne M, Lindgren M, Roll M, and von Koss Torkildsen J. Lund: Lund University, Media Tryck, 2009, p. 193-206.

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

Desmond JE, Gabrieli JDE, and Glover GH. Dissociation of frontal and cerebellar activity in a cognitive task: evidence for a distinction between selection and research. *Neuroimage* 7: 368-376, 1998.

Doya K. What are the computations of the cerebellum, the basal ganglia and the cerebral cortex? *Neural Networks* 12: 961-974, 1999.

Drepper J, Timmann D, Kolb FP, and Diener HC. Non-motor associative learning in patients with isolated degenerative cerebellar disease. *Brain* 122: 87-97, 1999.

Ferrand L and New B. Semantic and associative priming in the mental lexicon. In: *Mental lexicon: Some words to talk about words,* edited by Bonin P. Hauppauge, New York: Nova Science, 2003, p. 25-43.

Ferrucci R, Marceglia S, Vergari M, Cogiamanian F, Mrakic-Sposta S, Mameli F, Zago S, Barbieri S, and Priori A. Cerebellar transcranial direct current stimulation impairs the practice-dependent proficiency increase in working memory. *Journal of cognitive neuroscience* 20(9): 1687- 1697, 2008.

Fiez JA, Petersen SE, Cheney MK, and Raichle ME. Impaired non-motor learning and error detection associated with cerebellar damage. A single case study, *Brain* 115: 155–178, 1992.

Forster KI and Davis C. The density constraint on form-priming in the naming task: Interference from a masked prime. *Journal of Memory and Language* 30: 1–25, 1991.

Huang Y, Edwards MJ, Rounis E, Bhatia KP, and Rothwell JC. Theta burst stimulation of the human motor cortex. *Neuron* 45: 201-206, 2005.

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

Ichimiya T, Okubo Y, Suhara T, and Sudo Y. Reduced Volume of the Cerebellar Vermis in Neuroleptic-Naive Schizophrenia. *Biological Psychiatry* 49: 20–27, 2001.

Jarvis BG. DirectRT (Version 2008.1.0.11) [Computer software]. New York: Empirasoft Corporation, 2008.

Kiss GR, Armstrong C, Milroy R, and Piper J. An associative thesaurus of English and its computer analysis. In: *The Computer and Literary Studies,* edited by Aitken AJ, Bailey RW, and Hamilton-Smith N. Edinburgh: Edinburgh University Press, 1973, p. 153-165.

Koch G, Mori F, Marconi B, Codecà C, Pecchioli C, Salerno S, Torriero S, Lo Gerfo E, Mir P, and Oliveri M. Changes in intracortical circuits of the human motor cortex following theta burst stimulation of the lateral cerebellum. *Clinical Neurophysiology* 119: 2559 –2569, 2008.

Larsell O and Janssen J. *The comparative anatomy and histology of the cerebellum, Vol 3 III: The human cerebellum, cerebellar connections and cerebellar cortex*. Minneapolis: University of Minnesota Press, 1972.

Lucas M. Semantic priming without association: a metaanalytic review. *Psychonomic Bulletin and Review* 1: 618– 630, 2000.

Martin P and Albers M. Cerebellum and schizophrenia: a selective review. *Schizophrenia bulletin* 21(2): 241-250, 1995.

McEvoy CL and Nelson DL. Category name and instance norms for 106 categories of various sizes. *American Journal of Psychology* 95(4): 581-634, 1982.

Miall RC and Christensen LO. The effect of rTMS over the cerebellum in normal human volunteers on peg-board movement performance. *Neuroscience Letters* 371: 185–189, 2004.

Miall RC and King D. State estimation in the cerebellum. *Cerebellum*, 7, 572–576, 2008.

Molinari M, Filippini V, and Leggio MG. Neuronal plasticity of interrelated cerebellar and cortical networks. *Neuroscience* 111(4): 863-870, 2002.

Moritz S, Mersmann K, Kloss M, Jacobsen D, Andresen B, Krausz M, Pawlik K, and Naber D. Enhanced semantic priming in thought-disordered schizophrenic patients using a word pronunciation task, *Schizophrenia Research* 48(2): 301-305, 2001.

Moss HE, Ostrin RK, Tyler LK, and Marslen-Wilson WD. Accessing different types of lexical semantic information: evidence from priming. *Journal of experimental psychology: Learning, Memory, and Cognition* 21(4): 863-883, 1995.

Nestor P, Valdman O, Niznikiewicz M, Spencer K, McCarley R, and Shenton M. Word priming in schizophrenia: Associational and semantic influences. *Schizophrenia Research* 82(2): 139-142, 2006.

Ober B. RT and non-RT methodology for semantic priming research with Alzheimer's Disease Patients: A critical review. *Journal of clinical and experimental neuropsychology* 24(7): 883-911, 2002.

Ohtsuka K and Enoki T. Transcranial magnetic stimulation over the posterior cerebellum during smooth pursuit eye movements in man. *Brain* 121: 429–435, 1998.

Okugawa G, Sedvall GC, and Agartz I. Smaller Cerebellar Vermis But Not Hemisphere Volumes in Patients With Chronic Schizophrenia. *American Journal of Psychiatry* 160: 1614–1617, 2003.

Okugawa G, Nobuhara K, Takase K, and Kinoshita T. Cerebellar Posterior Superior Vermis and Cognitive Cluster Scores in Drug-Naïve Patients with First-Episode Schizophrenia. *Neuropsychobiology* 56: 216–219, 2007.

Oliveri M, Koch G, Torriero S, and Caltagirone C. 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, 2005.

Paulin MG. A Kalman filter theory of the cerebellum. In: *Dynamic Interactions in Neural Networks: Models and Data*, edited by Arbib MA and Amari SI. New York: Springer, 1989, p. 239-259.

Pedersen T, Patwardhan S, and Michelizzi J. Wordnet::similarity - measuring the relatedness of concepts. *Proceedings of the Nineteenth National Conference on Artificial Intelligence (AAAI-04)*, 1024–1025, 2004.

Perea M and Gotor A. Associative and semantic priming effects occur at very short stimulus-onset asynchronies in lexical decision and naming. *Cognition* 62: 223-240, 1997.

Perea M and Rosa E. The effects of associative and semantic priming in the lexical decision task. *Psychological Research* 66: 180–194, 2002.

Picard H, Amado I, Mouchet-Mages S, Olié J-P, and Krebs M-O. The Role of the Cerebellum in Schizophrenia: an Update of Clinical, Cognitive, and Functional Evidences. *Schizophrenia Bulletin* 34(1): 155–172, 2008.

Pickering MJ and Garrod S. Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences* 11(3): 105-110, 2007.

Postman L and Keppel G. *Norms of word associations*. New York: Academic Press, 1970.

Rogers SL and Friedman RB. The underlying mechanisms of semantic memory in Alzheimer's disease and semantic dementia. *Neuropsychologia*, 46: 12-21, 2008.

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

Theoret H, Haque J, and Pascual-Leone A. Increased variability of paced finger tapping accuracy following repetitive magnetic stimulation of the cerebellum in humans. *Neuroscience Letters* 306: 29–32, 2001.

Thompson-Schill SL, Swick D, Farah MJ, D'Esposito M, Kan IP, and Knight RT. Verb generation in patients with focal frontal lesions: A neuropsychological test of neuroimaging findings. *Proceedings of the National Academy of Sciences of the Unites States of America* 95(26): 15855-15860, 1998.

Torriero S, Oliveri M, Koch G, Lo Gerfo E, Salerno S, Petrosini L, and Caltagirone C. Cortical networks of procedural learning: evidence from cerebellar damage. *Neuropsychologia* 45(6): 1208–14, 2007.

Van Overschelde JP, Rawson KA, and Dunlosky J. Category norms: An updated and expanded version of the Battig and Montague (1969) norms. *Journal of memory and language* 50: 289-335, 2004.

Wolpert DM, Miall RC, and Kawato M. Internal models in the cerebellum. *Trends in Cognitive Sciences, 2*(9), 338-347, 1998.

Woodruff-Pak DS, Papka M, and Ivry RB. Cerebellar involvement in eyeblink classical conditioning in humans. *Neuropsychology* 10: 443–58, 1996.

Hurford, J., Flaherty, M., & Argyropoulos, G. (2007). Past and future, human and non-human, semantic/procedural and episodic. *Behavioral and Brain Sciences*, *30*(3), 324-325. Commentary on T. Suddendorf & M. Corballis, *The evolution of foresight: What is mental time travel and is it unique to humans?*

Past and future, human and nonhuman, semantic/procedural and episodic

DOI: 10.1017/S0140525X07002105

James R. Hurford, Molly Flaherty, and Giorgis Argyropoulos

Language Evolution and Computation Research Unit, School of Philosophy, Psychology and Language Sciences, University of Edinburgh, Edinburgh, EH8 9LL, United Kingdom.

jim@ling.ed.ac.uk **http://www.ling.ed.ac.uk**
s0672076@ling.ed.ac.uk **s0677134@ling.ed.ac.uk**

Abstract: The overlap of representations of past and future is not a completely new idea. Suddendorf & Corballis (S&C) usefully discuss the problems of testing the existence of such representations. Our taxonomy of memory differs from theirs, emphasizing the late evolutionary emergence of notions of time in memory.

The target article makes a useful contribution. We offer some reservations that do not undermine its central purpose.

Suddendorf & Corballis (S&C) join company with the literature from the past three decades casting doubt on a radical psychological past/future (memory/planning) distinction. Whereas memory was once thought of as exclusively retrospective, it is now recognized that "prospective memory" (a term introduced by Meacham & Singer 1977) shares features with retrospective memory. Cook et al. (1983) also demonstrated the overlap between retrospective memory and prospective memory in rats. Tulving (2005) identifies episodic memory closely with both past and future (i.e., planned) events.

We are not so convinced that nonhumans lack any retrospective episodic memory. Sue Savage-Rumbaugh (personal communication) tells of promising Kanzi, the well-known bonobo, a treat tomorrow and then being reminded of the promise the next day. Kanzi's specific desire may not have waned overnight, so there was not necessarily any representation of a future mental state distinct from the present, but Kanzi did recall the specific event of being promised. For retrospective cases, the question is whether, when an animal observes something happening, it stores a stripped-down "episodic" memory of the particular event, or makes some inferences about the lasting state of the world, such as "food behind tree." King, the gorilla tested by Schwartz et al. (2004), was shown events with no lasting consequence on the world, such as a man skipping. King could show that he recalled these events, as compared with distractor suggestions of other events which he might have (but had not) seen. But King was only tested 15 minutes after observing the event. The data are sparse, and we agree with S&C that there is a large quantitative gulf between humans and nonhumans for retrospective episodic memory.

Designing experiments to test the distinction concerned is challenging. We suggest a sense in which considerations of parsimony can conceivably be applied. Obviously no animal, human or otherwise, stores all the information from an observed event. If an animal remembers a specific episode, how much selection of its details happens? If, on the other hand, the animal does not store the episodic information, but only certain inferences relevant to its own typical needs, how many such inferences does it make? S&C's itemization of the components of "www" is useful. Hurley (2006) describes humans as "inferentially promiscuous." Does there come a point when we have to list so many different types of "w" inferred by an animal from an event that it becomes more parsimonious to assume that the animal just stored a stripped-down episodic memory of the event itself?

We agree with S&C that Mulcahy and Call (2006) come close to demonstrating planning by animals for anticipated needs. Mulcahy and Call (2006) showed that some bonobos in their experiment collected a specific tool for use on a specific food-retrieving task as much as 14 hours later. S&C's scrupulous critique of this experiment is correct, and there indeed remains the possibility that the animals just got into the habit of selecting the right tool. A very similar procedure to S&C's suggested red and green light future-planning food caching experiment has in fact been successfully completed by scrub jays (Raby et al. 2007), but S&C would judge that this is also not sufficient, because of the limited scope and inflexibility of the behavior. If these objections are valid, then radically different criteria for thinking about mental time travel (MTT), far from any sort of www components, are in order. We commend S&C for setting out just such criteria.

S&C's theater metaphor is a bold departure, and it provides fruitful ground for new investigation. That said, several components show evidence of insufficient development. Most particularly, it is not clear that the "broadcaster" component is integral to their framework. Does episodic memory only become episodic when it is shared with others? It seems unlikely that S&C would wish to make this claim, but by including the broadcaster as one of their seven components, the reader is left to think that episodic memory requires verbal sharing. This precludes possible research on episodic, or episodic-like, memory in nonhumans.

Although it plays no central role in S&C's main discussion of MTT, we take issue with their taxonomy of memory and prospection systems in their Figure 1. We see the picture as radically different. We would bracket procedural and semantic memory together as timeless forms of memory, evolutionarily preceding the emergence of episodic memory, in both its retrospective and prospective forms. Indeed, when S&C appear to distinguish between "semantic memory and prospection," we find this hard to interpret. Semantic memory encodes tenseless facts like "sugar be sweet." Likewise, procedural memory encodes instructions on what action to take whenever certain circumstances arise. Both procedural and semantic memories are laid down by experience, which is necessarily temporally prior to their formation, but they are not memories about particular events in the past, still less about events in the future. So we agree with S&C's statement that "The mental reconstruction of past events and construction of future ones may have been responsible for the concept of time itself, and the understanding of a continuity between past and future" (sect. 2.1, para. 3). This entails that the emergence of past/future episodic memory was also the emergence, for the first time in evolution, of types of mental representation incorporating notions of past and future time. This is, of course, a different matter from semantic or procedural memories encoding facts about temporal order, for example, that thunder follows lightning.

There is now increasing evidence for coinvolvement of motor and sensory components in memory for objects and actions, tending to conflate the semantic/procedural distinction. Hommel et al. (2001) give an overview pointing to the conclusion that "Perceived events (perceptions) and to-be-produced events (actions) are equally represented by integrated, task-tuned networks of feature codes – cognitive structures we call event codes" (p. 849). Hurford (2007) gives a short survey of other evidence; see also Martin et al. (1996), who found that "naming tools selectively activated a left premotor area also activated by imagined hand movements." Similarly, Mecklinger et al. (2002) concluded that "visual working memory for manipulable objects is based on motor programmes associated with their use" (p. 1115). Gibson's (1979) idea of affordances depends crucially on perception-action coupling. The discovery of mirror neurons also shows overlaps between sensory and motor representations of actions. In philosophy, a school of thought known as "Enactive Perception" also emphasizes the involvement of action in perception; see Noë (2004).

Finally and quite speculatively, we mention a kind of mental representation which, though essentially semantic/procedural, encodes information about typical types of complex events brought to mind by humans in either a retrospective or a prospective sense. This is the idea of "scripts," such as the complex structure of things to do when going to a restaurant (Schank & Abelson 1977). Such scripts form a framework for our retrospective and prospective representations of events. The ability to conjure up such scripts is a kind of general mental space integration capacity.

Memory, imagination, and the asymmetry between past and future

DOI: 10.1017/S0140525X07002117

Bjorn Merker
Gamla Kyrkvagen 44, SE-14171 Segeltorp, Sweden.
gyr694c@tninet.se

Abstract: A number of difficulties encumber the Suddendorf & Corballis (S&C) proposal regarding mental time travel into the future. Among these are conceptual issues turning on the inherent asymmetry of time