Argyropoulos, G. P. (2008). The subcortical foundations of grammaticalization. In Andrew D. M. Smith, Kenny Smith and Ramon Ferrer i Cancho (eds.), The Evolution of Language: Proceedings of the 7th International Conference on the Evolution of Language, Singapore:World Scientific Press, pp. 10-17.

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.

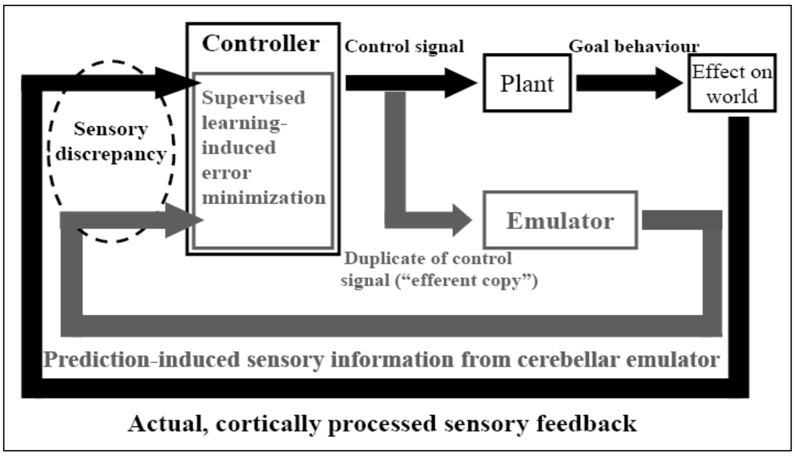

Figure 1. A simple cerebellar feedforward emulator (e.g., Wolpert, Miall, & Kawato, 1998). The predictions of the internal model are constantly updated, based on the error signals of the discrepancy induced by the actual sensory feedback.

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).

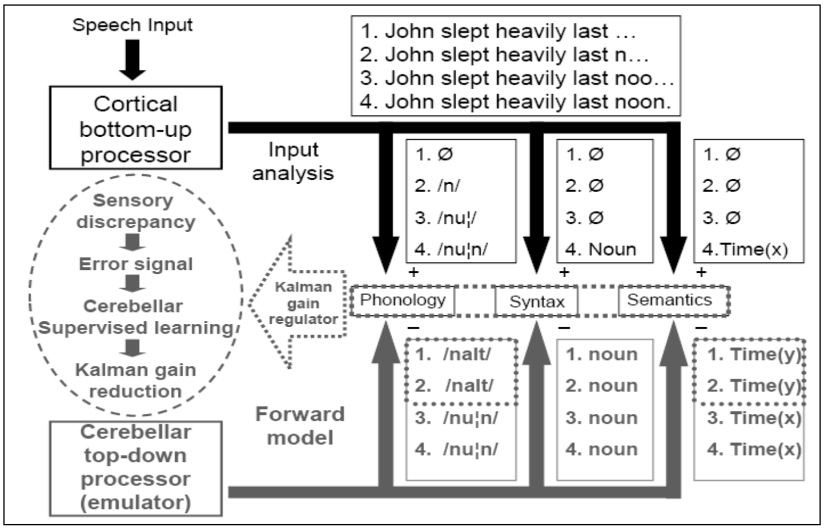

Figure 2. The CB as a domain general Kalman Filter (Paulin 1989) meets Pickering and Garrod (2007): Minimization of the gain of the Kalman regulator via routinization corroborates the reliance on a top-down (expectation-based) processing modality; "shallow processing" (Barton & Sanford, 1993) and "good-enough 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).

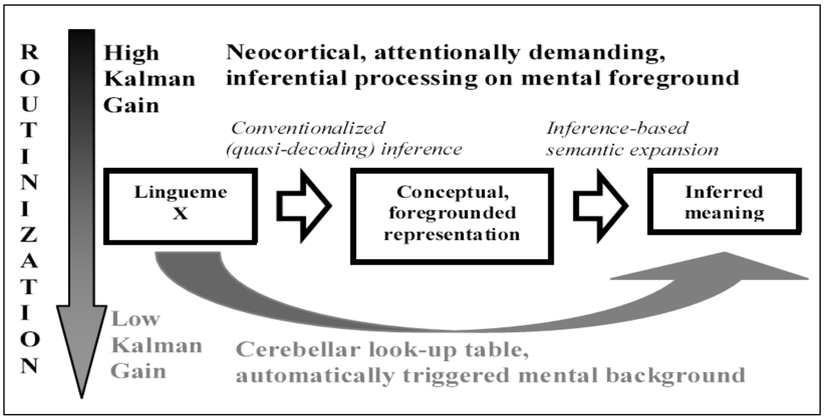

Figure 3. Routinization-induced proceduralization of conceptual encoding meets the "cerebellarization" of cognitive repertoires.

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)**.

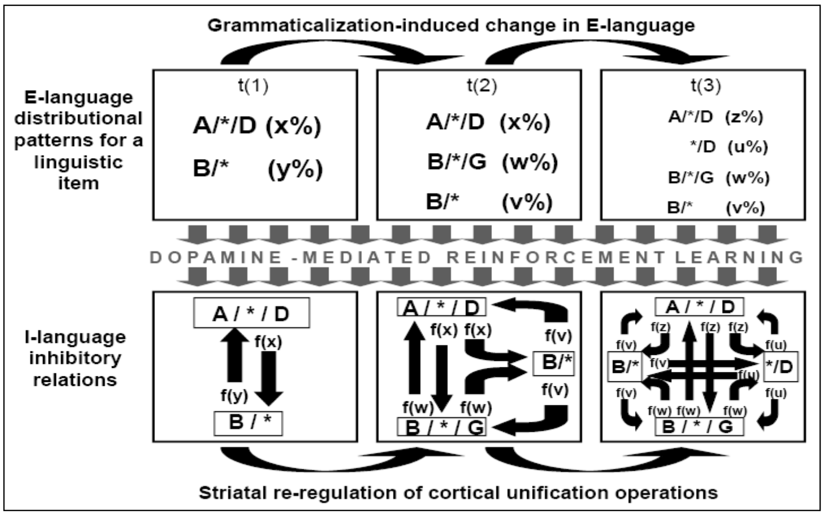

Figure 4. **t(1)-t(2):** a member of category G frequently co-occurs with the sequence B/*/, which triggers the strengthening of its probabilistic representation in the frontostriatal circuit, and thus the strength of the inhibitory signals sent to the competing alternatives: a gradual "obligatorification" (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.