

Can DPM Brain Agents Write Stories and Sing Songs?

James C Houk

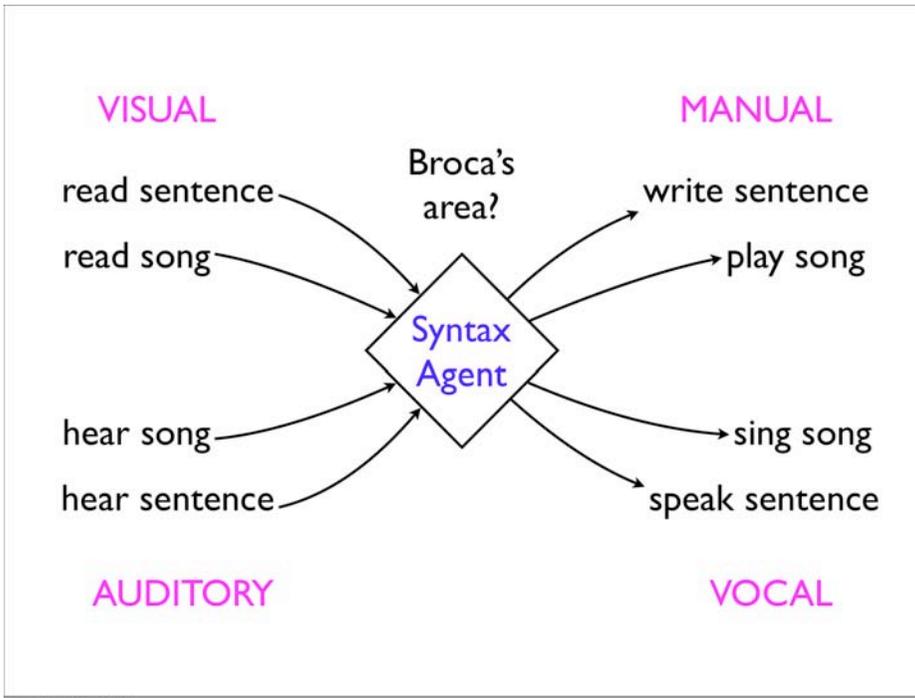
Northwestern University
Department of Physiology
j-houk@northwestern.edu

In the introduction to this workshop on Syntax in the Brain [Houk, 2011 (this volume)], I suggested that the concept of syntax might apply to movement, as well as to language and music. According to this view, the motor system has to discover different *motor syntaxes*, ones that specify viable sequences with which movements (and submovements) must be executed in order to achieve their goals. Subsequent speakers picked up on this theme, and we came to recognize that motor syntax, which suffers a disorder in schizophrenia, is a prominent feature of movement control, quite analogous to language syntax being a prominent feature of speech production. Now I would like to emphasize reasoning in the opposite direction – what can language, music, and other higher cognitive functions, learn from the neuroscience of motor systems.

In his book *Music, Language, and the Brain*, Patel [2008] did a remarkable job of elaborating on the many syntactic parallels between music and language, at the level of psychology and rule-based operations. Throughout this workshop on syntax in the brain, we have not pursued these issues very much. Instead, we have followed a neuroscientific perspective, which is illustrated schematically in Figure 1. This diagram shows how a Syntax Agent, defined as a neuroscientifically constrained brain module, might serve as a “common resource” for syntax, whether it be for movement, for music, for language, or perhaps even for other high-level cognitive functions such as mathematics [Prado et al 2011].

1 Syntax Agent

What I am seeking is an enunciation of what a generic syntax agent might be, coming at it from a neuroscientific perspective. The Simple Recurrent Networks (SRNs) described in my introductory talk are promising candidates, but they have important limitations, some of which were discussed in Tabor [2011 (this volume)] and in Ohta et al [2011 (this volume)]. Instead of SRNs, might the brain's generic syntax agents be the Distributed Processing Modules (DPMs) that I also described in my introductory talk [Houk 2011 (this volume)]? If so, we might profit from the vast neuroscientific knowledge we now possess, and continue to accumulate, about the control of sequential movements.



Tuesday, May 17, 2011

Figure 1: Schematic of a neuroscientifically constrained Syntax Agent that formulates MANUAL and VOCAL motor plans for both music and language based on VISUAL and AUDITORY modalities of input from association areas of the brain. Broca's area is highlighted, because patients have deficits in musical syntax in addition to deficits in language [Patel 2005].

In order to fathom a DPM's capacity for learning and implementing syntactic operations, more needs to be said about DPM composition [Houk 2011]. *Each DPM is a mesoscopic module*, comprised of a large array of *microscopic* modules. Both scales of modularity utilize the same impressive set of learning and processing operations,

designated the DPM architecture in Figure 2, which is discussed later. In any case, when thinking about the hypothetical Syntax Agent in Figure 1, consider it to be an array of microscopic DPM modules. The whole array receives massive visual and auditory input that is related to both sentences and songs. Its vocal outputs are plans for controlling speech and song, whereas its manual outputs are plans for writing text and playing musical instruments. What a spectacular common resource for learning and implementing syntax! At a yet higher *macoscopic* level, there must be an array of DPMs for processing phonemes, words, and sentences, in order to explain language. This idea is elaborated in section 2. For the moment, let's continue thinking of just one syntax agent, a mesoscopic DPM.

The different kinds of syntax present in music, or in language, could be learned more or less independently by different sets of microscopic DPM modules that comprise the hypothesized Syntax Agent in Figure 1. Therefore, the fundamentally different rule-based operations conceived of by linguists, versus those conceived of by musicians, should present no problem for this neuroscientifically-based syntax agent.

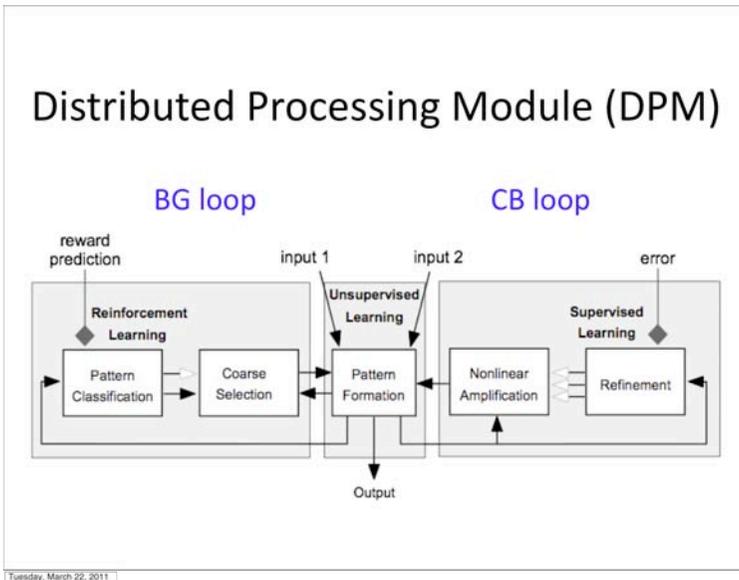


Figure 2: Schematic of the DPM architecture. It combines three distinct kinds of processing with three different kinds of learning, unsupervised learning in cerebral cortex, reinforcement learning in basal ganglia, and supervised learning in cerebellum.

What makes the DPM architecture such a powerful learning and processing agent? In a nutshell, this is because it profits from the simultaneous actions of three qualitatively different kinds of signal processing operations, each with its own specialized learning rule [Houk 2010]. Figure 2 illustrates these ideas schematically. The cerebral cortex operates much like the classical idea of a neural network. It learns from prac-

tice (unsupervised Hebbian learning) how to combine several input vectors in order to form useful outputs. The loops through basal ganglia and cerebellum tell it what to practice. How do these subcortical loops know what to practice? The loop through basal ganglia uses a sophisticated pattern classification step to evaluate, through reinforcement learning, what is currently going on up in the area of cerebral cortex that projects to it. The immediate consequence of this evaluation is a coarse selection of good activity and de-selection of bad activity. The loop through cerebellum amplifies and refines what the basal ganglia loop selects as salient, and if an error is made, the cerebellar cortex learns not to make the same mistake in the future, using a simplified version of supervised learning. The guidance that is provided by these two subcortical loops teaches the cerebral cortex to generate appropriate outputs more rapidly and more accurately – they become habits.

A key feature of the DPM architecture is its capacity to use all three kinds of learning: unsupervised learning in cortex, reinforcement learning in basal ganglia, and supervised learning in cerebellum. The SRN architecture is not constrained by a biologically feasible learning rule. Instead, the network is trained in any way possible, frequently by using back-propagation over time. Nevertheless, difficulties in learning are encountered, due to the operational limits of attractor networks [Tabor 2011 (this volume)].

Given the generic nature of the operations performed by a DPM, how can different DPMs perform different functions and implement different kinds of rules? This happens not only during trial-to-trial learning, but also as a consequence of developmental plasticity. In the adult brain, each mesoscopic DPM receives different combinations of input to operate upon. For example, the DPM that subserves the motor cortex receives a variety of sensory inputs and sends outputs that command activations of diverse muscles, thus performing a sophisticated and multifaceted sensory-motor function. In contrast, the Syntax Agent in Figure 1 receives processed auditory and visual inputs that represent songs and words, and it has outputs that represent plans for hand and vocal actions. In this manner, a generic DPM signal processing agent could subserve a variety of cognitive and sensorimotor functions. In the next section I discuss the manner in which arrays of mesoscopic DPMs can be linked together in order to perform tasks that require both cognitive AND sensorimotor functions?

2 Arrays of mesoscopic DPM agents

I will try to explain how DPMs might be used in combination to implement a set of functions that are required to perform an entire task, like creating a thought in your mind and then writing a sentence about it. The obvious approach is to try using arrays of mesoscopic DPM agents that communicate with each other. Let's begin with a more manageable task, namely the Replicate task that was described in the earlier workshop paper on schizophrenia [Fraser & Houk, 2011 (this volume)]. As implied there, a complete model of Replicate would require a minimum of three DPMs. One

would be needed to encode a sequence of visual targets into a serial order working memory. In psychological terms, its output is like a ‘thought’ about an action sequence. Its output vector is a representation of that thought. A second DPM would be needed to decode the thought into a sequence of actions. Its output vector could be called a ‘plan’ in psychological terms, and that DPM’s output vector is a representation of the plan. A third DPM is required to generate the set of motor commands that are sent to the brainstem or spinal cord to execute any given action, be it vocalization

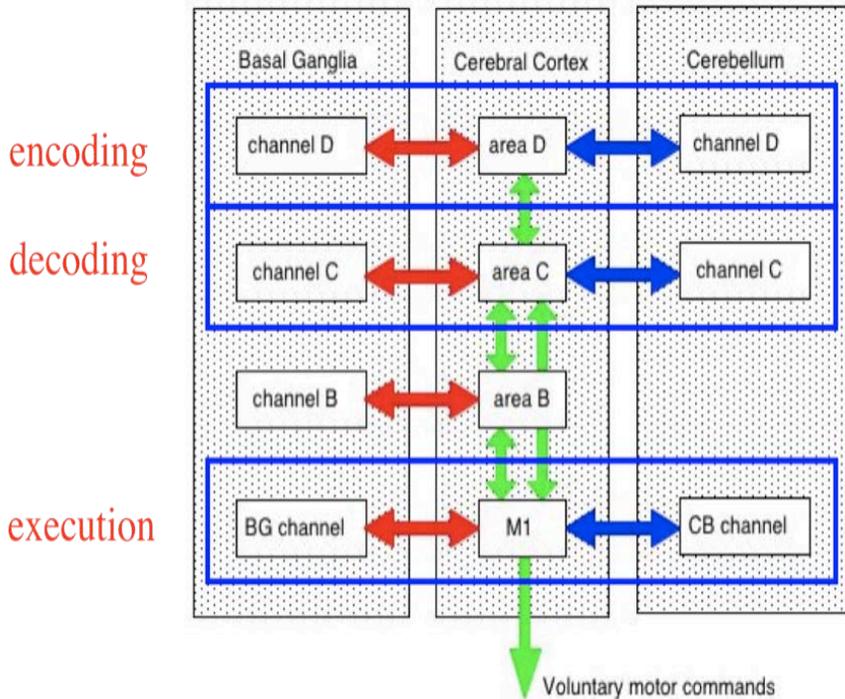


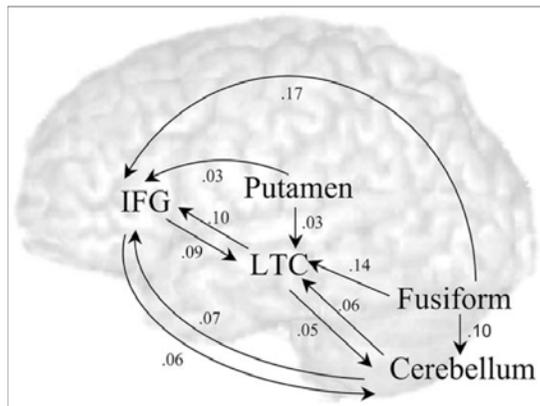
Figure 3: A network of DPMs that could potentially perform the three key operations required for the Replicate task, namely encoding, decoding and execution. The diagram labels 4 areas of cerebral cortex. The 3 contained within the 3 blue rectangles are presumed to perform the operations of encoding, decoding, and execution; the other one is presumed to be inactive in this example. The cortical areas are labeled M1, B, C and D, where M1 = primary motor cortex, B & C represent premotor areas, and D represents the prefrontal working memory area DLPFC discussed by Fraser & Houk [2011 (this volume)]. Loops through the basal ganglia are illustrated schematically with bidirectional red arrows, and loops through cerebellum are labeled with bidirectional blue arrows. Thus each blue rectangle represents a DPM, a generic agent whose computational steps are summarized in Figure 2. The bidirectional green arrows represent cortico-cortical pathways, which are predominantly excitatory. They represent the main channels for communication among DPMs. (Note that the green arrow connecting M1 with area C actually passes underneath area B, no connections intended.)

or be it hand movements. Its output vector can be called a ‘command’ in psychological terms. Much is known about the representation of voluntary motor commands for hand movements in M1, the primary motor cortex [Georgopoulos 1995].

3 Testing the theory

How might these ideas be tested? Functional imaging of the brain is an obvious choice, although it is limited by spatial and temporal resolution. One approach that has been applied to address the temporal resolution problem is dynamic causal modeling [Friston et al 2003]. We recently applied this method to a word processing task [Booth et al 2007]. We examined the role of the cerebellum and basal ganglia during a rhyming task that was presented in the visual modality. We used Dynamic Causal Modeling (DCM) to look at the directional influence of one brain region on another. We found that the cerebellum had reciprocal modulatory influences and the basal ganglia had unidirectional modulatory influences with brain regions involved in phonological processing, as illustrated in Figure 4.

Figure 4: Modulatory effects of the rhyming task on the language network. Averaged strengths of the effects across individuals are indicated. IFG = left inferior/middle frontal gyri, LTC = left superior/middle temporal gyri.



The cerebellum had reciprocal connectivity with two cortical areas involved in phonological processing, the IFG and the LTC, but the putamen of the basal ganglia had only unidirectional connectivity with these same two phonological areas of cortex. This supported the suggestion that the basal ganglia may be involved in the cortical initiation of coarse or “ball-park” phonological representations in these structures, and the cerebellum was involved in the amplification and refinement of these representations. It is important to clarify that these results do not show a lack of input from cortex to putamen during the task, because intrinsic connections were significant. These connections just were not differentially modulated in the rhyming task compared to the control task (line judgement). This finding seems appropriate given that the putamen, in the DPM theory, is involved in pattern classification of essentially equivalent activations from cortex in both tasks. ***Most importantly, these findings***

support the hypothesis that loops through basal ganglia and cerebellum are important components of higher cognitive function.

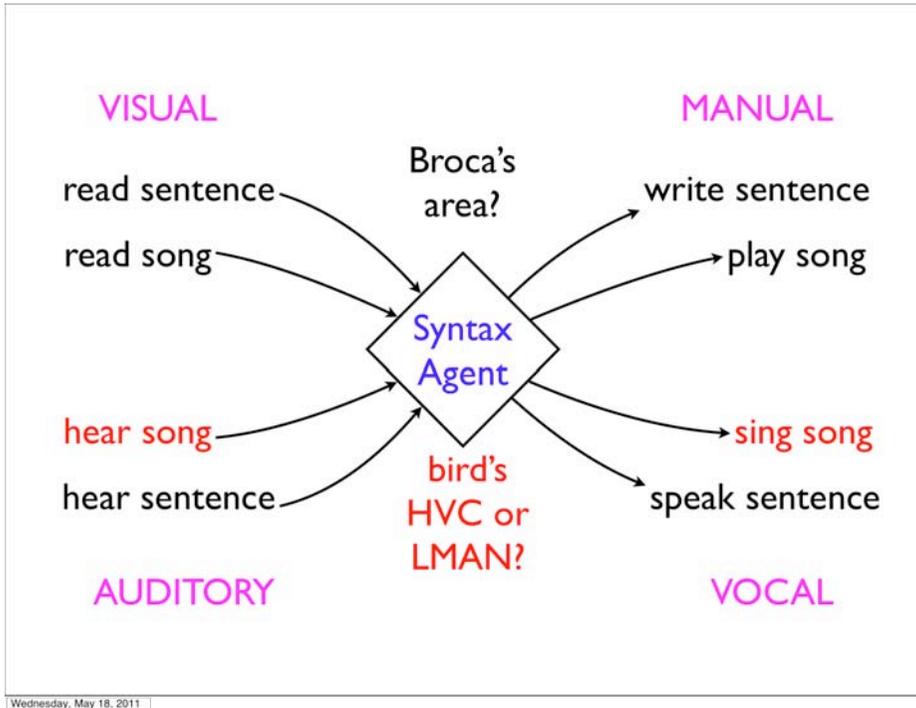


Figure 5: The schematic of a Syntax Agent in Figure 1 is highlighted in red to illustrate the birdsong model of language. HVC and LMAN are two regions of the songbird's pallium that are critical for birdsong.

Functional imaging also suffers from spatial resolution problems, being limited to approximately the mesoscopic scale of DPMs. Microelectrodes are needed in order to monitor signal processing in the microscopic modules. This powerful recording methodology can most readily be accomplished in animal models. In recent years advantageous use has been made of the songbird model of language for testing ideas about circuit functionality [Doupe et al 2005]. What is currently lacking in the songbird model is a reliable relationship between specific regions of songbird pallium and specific areas of cerebral cortex in primates. Figure 5 relates the Syntax Agent concept to birdsong by highlighting in red the pathway between hearing a song and singing a song. It also asks whether Broca's area might be analogous to either HVC or LMAN, two regions of the songbird's pallium that are critical for birdsong. The finding that songbirds can in fact do complex syntactic processing [Gentner et al 2006] suggests that these questions are worth pursuing.

Another approach for testing these ideas is through computer simulation. Within the framework of agent-based modeling [Wilensky & Rand 2011], each mesoscopic DPM can be thought of as an individual agent. This is useful because the cerebral network as a whole can then be considered as a multi-agent system. Each DPM receives different inputs and thus computes different outputs, even though the information processing operations of all of the agents is essentially the same. Agent-based modeling could then be used to systematically test the collective effects of individual DPM operations [Houk 2011]. Given the novelty of each agent's set of signal processing operations, combined with the agent's unusual learning capacity, sets of such agents might allow the emergence of language, intelligent behavior, and superior social interaction. With an appropriate model, these ideas could be tested.

4 Potential enhancements of the DPM architecture

The current version of the DPM architecture uses a very simple model of the cerebral cortex. It would be interesting to enhance the cortical node by adding an SRN [Elman 2004]. A more ambitious endeavor would be to interface the subcortical loops with a more elaborate model of cortex, such as the one being pursued by Hawkins [2004].

Pattern classification in the striatum of the basal ganglia is a critical step in brain processing. Considerable attention needs to be devoted to the construction of an adequate model of this operation. Certainly we need to understand the role of presynaptic versus postsynaptic inhibition as a mechanism for mediating competition [Ohta & Gunji 2006; Ohta et al 2011 (this volume)]. Presynaptic inhibition is also important in attentional gating of activity in the striatum [Ashby & Crossley 2011; Ding et al 2010], and this needs to be combined with the competitive pattern classification operation in order to simulate meaningful arrays of DPM agents.

The brain's representation of space [Peterson 2011], implicit memory [Eichenbaum 2009] and time [McClelland 2011 (this volume)] were addressed in the last three talks in this workshop. How these topics all fit together with concepts of brain agents will require much more work. I would only urge that investigators pay heed to the reciprocal connections between the hippocampus and the cerebellum that were described some years ago by Newman and Reza [1979]. Might the hippocampus also profit from the DPM architecture?

Acknowledgement

Dr. Houk received support for this work from grant NINDS P01NS044393.

Bibliography

- [1] Ashby, F. G. and M. J. Crossley (2011). "A computational model of how cholinergic interneurons protect striatal-dependent learning." *Journal of Cognitive Neuroscience*(in press).
- [2] Booth, J. R., L. Wood, D. Lu, J. C. Houk and T. Bitan (2007). "The role of the basal ganglia and cerebellum in language processing." *Brain Res* 1133(1): 136-44.
- [3] Ding, J. B., J. N. Guzman, J. D. Peterson, J. A. Goldberg and D. J. Surmeier (2010). "Thalamic gating of corticostriatal signaling by cholinergic interneurons." *Neuron* 67(2): 294-307.
- [4] Doupe, A. J., D. J. Perkel, A. Reiner and E. A. Stern (2005). "Birdbrains could teach basal ganglia research a new song." *Trends Neurosci* 28(7): 353-63.
- [5] Elman, J. L. (2004). "An alternative view of the mental lexicon." *Trends Cogn Sci* 8(7): 301-6.
- [6] Eichenbaum, H. and N. J. Fortin (2009). "The neurobiology of memory based predictions." *Philos Trans R Soc Lond B Biol Sci* 364(1521): 1183-91.
- [7] Fraser, D. and J. C. Houk (2011). "Motor syntax disorder in schizophrenia." *Proceedings of the Eighth International Conference on Complex Systems* (this volume), NECSI, in press.
- [8] Friston, K. J., L. Harrison and W. Penny (2003). "Dynamic causal modelling." *Neuroimage* 19(4): 1273-302.
- [9] Gentner, T. Q., K. M. Fenn, D. Margoliash and H. C. Nusbaum (2006). "Recursive syntactic pattern learning by songbirds." *Nature* 440(7088): 1204-7.
- [10] Georgopoulos AP (1995) Motor cortex and cognitive processing. In MS Gazzaniga (Ed.) *The Cognitive Neurosciences*, pp 507-517, MIT Press, Cambridge, MA.
- [11] Hawkins, J. and S. Blakeslee (2004). *On Intelligence*, Times Books.
- [12] Houk, J. C. (2010). *Voluntary Movement: Control, Learning and Memory*. *Encyclopedia of Behavioral Neuroscience*. G. F. Koob, M. Le Moal and R. F. Thompson. Oxford, Academic Press. 3: 455-458.
- [13] Houk, J. C. (2011). Action selection and refinement in subcortical loops through basal ganglia and cerebellum. In *Modelling Natural Action Selection*. A. K. Seth, J. Bryson and T. J. Prescott, Cambridge University Press, Cambridge, in press.
- [14] Houk, J. C. (2011). "Syntax in the brain: Motor syntax agents." *Proceedings of the Eighth International Conference on Complex Systems* (this volume), NECSI, in press.

- [15] Newman, P. P. and H. Reza (1979). Functional relationships between the hippocampus and the cerebellum: an electrophysiological study of the cat. *J Physiol* 287: 405-26.
- [16] Ohta, H. and Y. P. Gunji (2006). "Recurrent neural network architecture with pre-synaptic inhibition for incremental learning." *Neural Netw* 19(8): 1106-19.
- [17] Ohta, H., Y. Nishida and J. C. Houk (2011). "Presynaptic inhibition and incremental learning in the striatum of the basal ganglia." *Proceedings of the Eighth International Conference on Complex Systems* (this volume), NECSI, in press.
- [18] Patel, A. D. (2005). "The relationship of music to the melody of speech and to syntactic processing disorders in aphasia." *Annals of the New York Academy of Sciences* 1060 (The Neurosciences and Music II: From Perception to Performance): 59-70.
- [19] Patel, A. D. (2008). *Music, Language, and the Brain*, Oxford University Press.
- [20] Peterson, B. W. (2011). "Pursuit of moving visual targets as a syntactic behavior." *Proceedings of the Eighth International Conference on Complex Systems* (this volume), NECSI, in press.
- [21] Prado, J., R. Mutreja, H. Zhang, R. Mehta, A. S. Desroches, J. E. Minas and J. R. Booth (2011). "Distinct representations of subtraction and multiplication in the neural systems for numerosity and language." *Human Brain Mapping*.
- [22] Tabor, W. (2011). "Recursion and recursion-like structure in ensembles of neural elements." *Proceedings of the Eighth International Conference on Complex Systems* (this volume), NECSI, in press.
- [23] Wilensky, U. and W. Rand (2011). *An introduction to agent-based modeling: Modeling natural, social and engineered complex systems with NetLogo*. Cambridge, MA, MIT Press.