

Syntax in the Brain: Motor Syntax Agents

James C. Houk
Northwestern University
j-houk@northwestern.edu

I conceived of this workshop after reading a review article by Ani Patel [Nature Neuroscience, 2003], entitled Language, Music, Syntax and the Brain. Patel summarized how the syntax of music and language follow different rules, which operate on both auditory and visual modalities of input, to produce distinct perceptions and plans, motor plans for many different kinds of action. Given that a plethora of mechanistic differences underpin each, it should be no surprise that different brain networks are activated when listening to music as opposed to listening to speech. However, both music and language obey analogous syntactical rules, and a few critical discoveries support the existence of common areas in the cerebral cortex that are used for both kinds of syntax. One interesting example is Broca's area [Patel 2005]. Reviewing the doctrine, Patel [2003] hypothesized that "common resources" are used to process both the syntax of music and the syntax of language. This concept promoted a vigorous debate, with many examples for and against contained in a new book, Language and Music as Cognitive Systems [P. Rebuschat et al 2011, Oxford University Press].

According to Patel's definition, "syntax is *a set of principles governing the combination of discrete structural elements.*" I thought to myself, by this generalized definition, syntax probably applies to movement as well as to language and music. Consider individual movements as the discrete structural elements. Typically we combine sequences of individual movements in order to accomplish a goal. Therefore, movement syntax should be the set of principles that govern sequential combinations of individual movements, in order to achieve a goal. The goals of movement are usually transparent, in analogy with the goals of most

sentences. [Hmm, we seem to have an interesting analogy here.] Let's consider an example.

When we are thirsty, we need to walk to the cupboard, fetch a glass, go to the sink, turn on the water faucet, fill the glass with water, turn off the faucet, bring the glass to our mouth, and finally drink – thus satisfying our thirst. If we alter the serial order in which we perform these actions, we may very well not relieve our thirst. Think about the many other sequences of actions that need to be executed, in an appropriate serial order, in order to satisfy our many goals in life. Serial order seems to be fundamental. The rules that determine satisfactory combinations of individual movements seem to be what is needed, to define the syntax of movement.

What about individual movements themselves? Does reaching to a single visual target also possess syntax? Our first speaker Neil Berthier has studied reaching behavior in infants. He will now talk to us about how infants begin reaching with a whole sequence of tiny inaccurate submovements, and they gradually progress to a large initial, and reasonably accurate, primary movement, with one or two corrective submovements that home in on the target. I asked Neil to discuss these data and the underlying syntax of reaching in infants. [\[Talk 1\]](#) The Syntax of Human Infant Reaching, Neil Berthier

Now I will present my own analysis of the syntax of submovements, based on data from adult humans and monkeys.

1 Syntax of submovements: the NPE model

Tangential velocity traces of hand movements in primate reaching tasks display multiple peaks, a phenomenon that has been well-known since [Woodworth's seminal paper in 1899]. However, the cause of these irregularities is still under debate. The traditional view has held that multiple peaks in velocity can be attributed to the use of overlapping submovements (OSMs) to correct reaching errors. More recent theories suggest that the same effect can be attributed to a continuous control system that is subject to delays and non-linearity [Bhushan & Shadmehr 1999; Kawato et al 1992; Shadmehr & Mussa-Ivaldi 1994; Sternad & Schaal 1999]. Whatever the underlying mechanisms, it can be difficult to detect the occurrences of OSMs in rapid movement trials. Several years ago we described a novel soft-symmetry algorithm that uses higher derivatives (velocity, acceleration, jerk & snap) of movement traces for detecting OSMs [Fishbach et al 2005]. Instead of insisting on perfectly symmetric trajectories, as in Novak et al [2000], soft-symmetry relaxes this constraint by a small degree. These algorithms allow single behavioral trials to be decomposed into their component parts, as illustrated in Figure 1. This permits the statistical analyses of submovement behavioral features.

Statistical analyses of behavioral features of rapid primary movements and their corrective submovements revealed a recurrent pattern in the organization of fast and accurate primate reaching [Fishbach et al 2007]. OSM onset times show appreciable variability and are highly correlated with the normalized amplitude of the primary movement. We hypothesized that this submovement initiation pattern is incompatible with a continuous model of movement control and that it fits a particular discrete model of movement control particularly well. According to this model, a correction is initiated when the normalized predicted error (NPE) exceeds a threshold value T .

Figure 1: Movement error correction by discrete submovements. This example shows the decomposition of a hand movement into a primary movement and an OSM (overlapping submovement). In the illustrated trial, the primary movement would fall short of the target. However, an OSM is generated by a predictive controller to correct the error. Note that the velocity trace shows an inflection and the acceleration (not shown) has a zero crossing. These features are detected by the soft-symmetry algorithm to identify the onset of the OSM. D is the target distance and L is the latency of the OSM. Modified from Fishbach et al [2007].

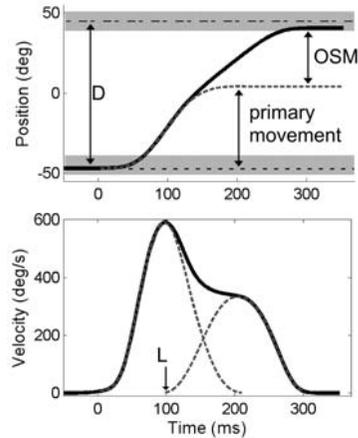


Figure 2 illustrates the NPE model. Note that predicted error applies to the ongoing trial, whereas prediction variance is computed over the subject's past performance with similar trials. These two factors, current error (also called *likelihood*) and prediction variance (related to the *prior*, as in prior knowledge), are the cornerstones of Bayesian models of decision-making [Koerding 2007].

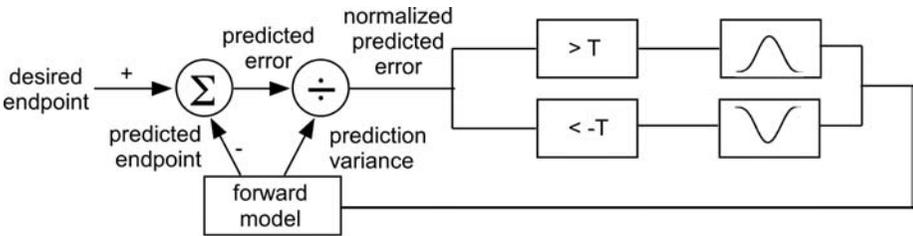


Figure 2: The NPE (Normalized Predictive Error) model of how corrective submovements are generated. Vision provides the information about the desired endpoint, which can be updated as rapidly as 180 ms when a visual perturbation is introduced at movement onset [Fishbach et al 2007]. The NPE model uses a type of forward model to compute the predicted endpoint based on both efference copy and sensory input, and it computes the prediction variance based on past experience. The normalized predicted error (Z-score) must exceed a threshold value T in order to initiate a corrective submovement. The

executed submovement follows an approximately bell-shaped velocity profile. Taken from Houk et al [2007].

We tested the NPE model against alternative control models using mechanistic implementations of simulated movements [Barringer et al 2008]. The output of each control model tested was fed into a fixed model of the neuromuscular system in order to generate statistical movement data for comparison between models and with monkey behavior. Amplitude-dependent noise was added to the pulse phase of the simulated command and no noise was added to the step phase. A simple mass-spring model with linear damping did not permit a clear distinction to be seen between a movement's endpoint, which exhibited variance as a consequence of motor command variance, and the system's equilibrium position, which showed no variance. Substituting fractional power damping for linear damping overcame that problem. Fractional power damping (proportional to velocity raised to the fifth power) has been used in previous modeling studies to capture the basic nonlinearity of the neuromuscular system [Barto et al 1999; Karniel & Inbar 1999; Houk et al 2002; Wang et al 2008]; it seems to be very important for effective damping in the control of rapid movements [Barto et al 1999] and explains the relationship between amplitude, velocity and duration of rapid human movements [Karniel & Inbar 1999].

Continuous control policies in general do not account for the large range of submovement onset times found in the animal data analyzed by Fishbach et al [2007]. In contrast, the use of a discrete control policy explained both the large range and the linear increase in submovement onset times as the normalized amplitude of the primary movement increased. All of the discrete control processes that we tested issued corrective movements that improved accuracy, and most of them shortened the duration of movement. However, the NPE policy outperformed, in terms of speed and accuracy [see Novak et al 2002 for a discussion of speed-accuracy tradeoff], a suite of other similar discrete policies for online error correction. Furthermore this policy resulted in trajectories that were qualitatively similar to primate reaching movements. Our findings are consistent with the hypothesis that both the NPE model and actual primate reaching behavior display performance that is essentially optimal. In essence, the NPE model computes the statistical significance of an error, and waits until that measure goes above a threshold value T to initiate a movement correction. ***The NPE model seems to characterize the syntax of individual movements quite well.***

Next let's consider a more sophisticated motor task. Paul Reber will talk about the Serial Interception Sequence Learning (SISL) task that he and his student Eric Gobel designed to study the representation of both time and space in the brain. I'll let Paul describe this task and relate its features to motor syntax. Paul will also discuss how statistical models of sequential movement relate to these data. [Talk 2] Models of sequential movements, Paul Reber.

How might one begin to think more neuroscientifically about syntactical rules, and how these rules characterize movement syntax? This issue troubled me, since I was aware of the huge controversy that pervades linguistics – there are so many theories of grammar that are widely disputed. Understanding the principles of motor syntax,

specifying the operations that the brain uses to plan and execute diverse motor tasks, might also be a complex problem.

Then I remembered Jeff Elman's 2004 elegant paper on an alternative view of the mental lexicon. Jeff, in fact, helped me initiate the planning of this workshop. He subsequently learned that he could not attend, so I decided I should summarize his paper for him.

2 Introduction to SRNs

A Simple Recurrent Network (SRN) is the agent used by Elman to learn the essence of the mental lexicon. Figure 3 shows the SRN. It is basically a 3-layer connectionist network, but in addition it has a recurrent pathway to Context units that store the output of the Hidden unit layer after each time step. At the next time step, the N hidden units receive a new stimulus from all of the Input units and a copy of the previous state which is stored in the Context units. The weights of all of these inputs are learned by training the network on a sentence prediction task. In the example I talk about here, each input unit represented a word, and the network learned to predict the next word in a set of practice sentences. The SRN was trained using back-propagation across time. After learning to perform at a high level, this network was then studied computationally in order to understand how it works, and to test its performance on some novel sentences that were formed by using an identical grammar.

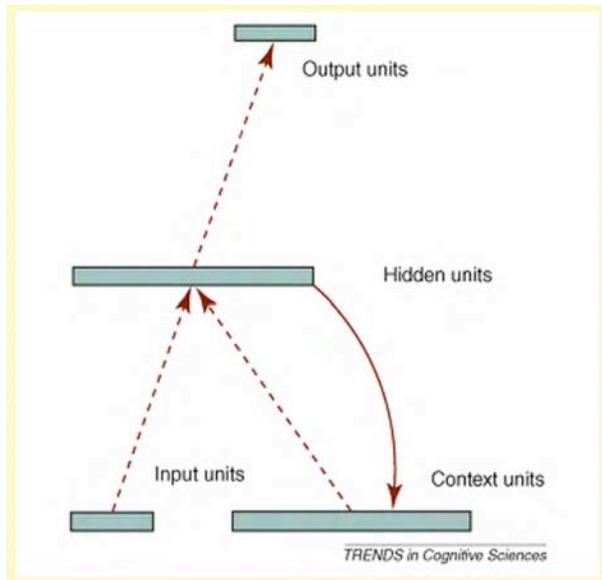


Figure 3: A Simple Recurrent Network. At each time step, the values of the Hidden units are copied to the Context units.

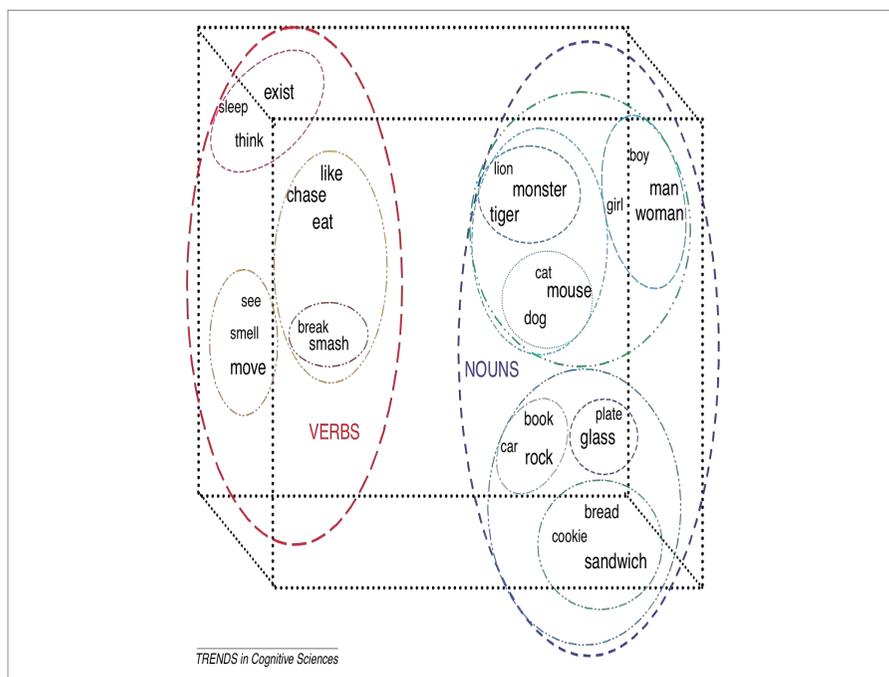


Figure 4: Schematic visualization of the high-dimensional state space after learning.

A schematic visualization, in 3D, of the high-dimensional state space described by the SRN's hidden-unit layer is shown in Figure 4. Through learning, the state space becomes naturally partitioned into different regions, regions that correspond to grammatical and semantic categories. The nesting relationship in space (e.g. man within Animate within Noun categories) reflect hierarchical relationships between categories.

I find these results quite remarkable. The N-dimensional activation of the hidden units convey the contents of the mental lexicon, and the weights of the synaptic inputs convey a neural equivalent of the lexicon itself.

I will now use another figure from Elman [2004] to illustrate an important issue in connectionist theory, the problem of catastrophic interference [McClelland et al 2010]. After teaching the SRN to perform at a high level, its performance was tested with a novel set of sentences, and the outcome is shown in Figure 5. This graph shows trajectories through 2 dimensions of the SRN's N-dimensional state space, while processing the verb 'runs' when it was preceded by different nouns. Differences in the location of the verb's state reflect systematic effects of context due to different preceding nouns. (PCA = principal component analysis.) The graph demonstrates that the SRN is sensitive to temporal context, a very important requirement for evaluating syntax.

Note that the 5 examples of 'runs' are located close to each other in this state space. To recognize these 5 examples as distinct entities requires a pattern classification operation in the output layer, in order to decide which pattern occurred. If the SRN had been trained to recognize the entire lexicon, the distinct entries would be much more densely clustered in state space. Superior pattern classification would then be needed to separately recognize the very large number of distinct patterns in the lexicon of any particular language. Think of the problems that would be faced by an SRN that had to learn to understand several different languages.

Our next speaker Whit Tabor will explain how recurrent pathways in the brain form attractor networks that help us understand how SRNs are able to do these recursion-like operations. Chomsky and colleagues [Hauser et al 2002] consider recursion to be a very powerful computational operation that underlies the rich, expressive, and open-ended power of human language. [Talk 3] Recursion and fractal grammars in attractor networks, Whitney Tabor.

Now, I would like to introduce my own approach to the quest for a generic brain agent, the one that was actually on my mind when I first conceived of this workshop.

3 Introduction to DPMs

The concept of Distributed Processing Modules (DPMs) as brain agents was originally deduced from the neuroanatomy and neurophysiology of the network that controls voluntary limb movement [Houk 2005]. The anatomical definition came out of the work by Peter Strick and his colleagues, using viral tracers to document trans-synaptic pathways [Kelley & Strick, 2000]. Each area of frontal cortex that has been investigated to date has an anatomical loop through the basal ganglia, and most of these areas also has an anatomical loop through the cerebellum [eg. Kelly & Strick

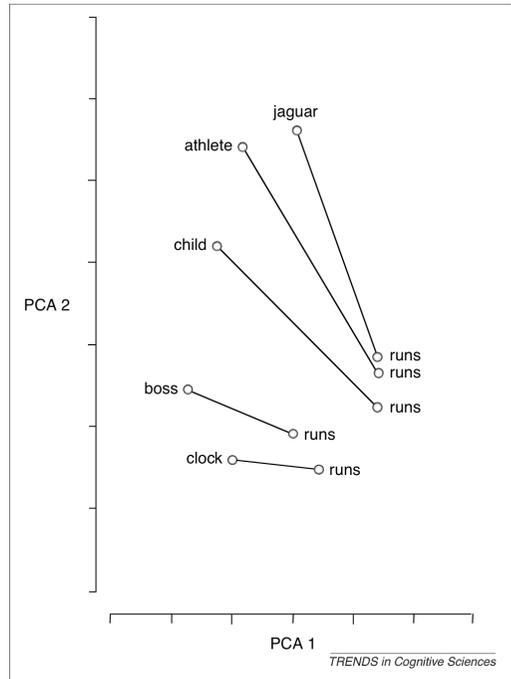


Figure 5: Trajectories through state space of the response 'runs' to five different preceding nouns.

2003, 2004]. I have worked on the physiology of these loops, and have come up with a summary assessment of the information processing operations that neural signals undergo when traversing these loops [Houk 2005]. The latest version of the DPM architecture is reproduced in Figure 6.

Each of approximately 100 cerebral cortical areas is individually regulated by relatively private loops through subcortical structures, particularly through the basal ganglia and cerebellum [Houk 2005; Houk & Wise 1995; Kelly & Strick 2003, 2004]. These DPMs have powerful computational architectures as summarized in Figure 6. Each DPM receives cortico-cortical input vectors from approximately 7 other DPMs (although only 2 are shown in Figure 6). The final outcome of all of the computations in a given DPM is a spatiotemporal pattern of activity in the module's output vector, representing the activity in its set of cortical output neurons. This output is sent as input to other DPMs, or to the brainstem or spinal cord. In this manner, arrays of DPMs form large-scale networks that function in combination to control behavior, or thought. The reader should consult Houk [2005] for a detailed description of this architecture and a justification of its capacity to control both actions and thoughts. The brief overview of functional operations in loops through basal ganglia (BG) and cerebellum (CB) given in the next two paragraphs is a summary that applies to the selection and initiation of movement commands that control discrete actions.

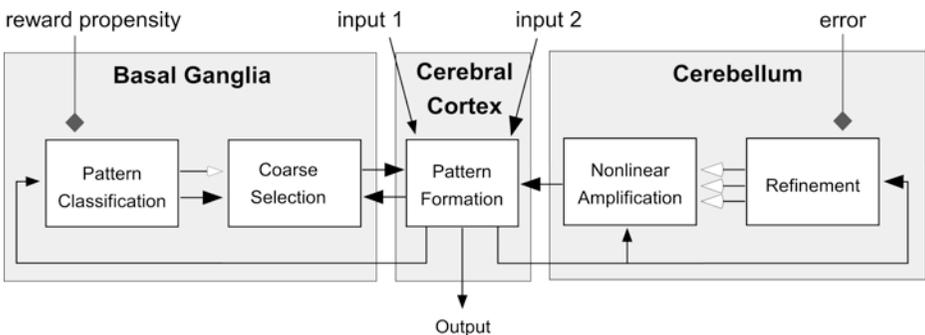


Figure 6: Schematic of a Distributed Processing Module (DPM). Relatively private loops through basal ganglia (BG) and cerebellum (CB) regulate the output vector transmitted by each module. In this manner, responses to input vectors by circuits within each area of cerebral cortex become elaborated by more sophisticated processing steps in BG and CB. Net excitatory pathways are shown with black arrowheads, net inhibitory pathways are shown with white arrowheads, and the diamonds signify neuromodulatory and training inputs. From Houk [2011].

A DPM's loop through BG is thought to regulate action selection [Gurney et al 2001; Houk 2001; 2005; Houk & Wise 1995; Redgrave et al 1999]. The label Coarse Selection is used in Figure 6 in recognition of the fact that BG output neurons recorded during a task comprised of multiple options typically do not select one

unique action. Instead they code for a small set of potential actions, for example forearm rotations with different amplitudes, velocities and directions of movement, as opposed to a unique movement [Gdowski et al 2007]. In spite of this qualification, coarse action selection poses a very difficult problem for a subject in a natural setting. While in my office, I have to decide whether I will pick up a writing instrument, type on my keyboard, scratch my chin, or walk to the kitchen to fetch a beverage (and think of all the other potential actions). On the input side of the BG loop in Figure 6, the Pattern Classification operation receives a huge vector of neural signals from one area of cerebral cortex (and functionally adjacent areas) plus a reward propensity signal from dopamine neurons in the midbrain. The cortical input is highly diverse in both qualitative and quantitative dimensions, providing a challenging computational problem for pattern classification. The reward signal produces a short-term attentional modulation of useful neuronal responses [Nicola et al 2000] plus a long-term consolidation of the synaptic weights that promoted them [Houk 2005]. In this manner, a diversity of events in a diversity of contexts can be classified with respect to their relevance and saliency. According to most contemporary models [reviewed in Houk 2007], bursts of medium spiny neurons in the striatum, via a direct pathway through BG, disinhibit their targets in thalamus, allowing thalamo-cortical loops to initiate tentative patterns of activity that represent coarse selections of appropriate actions. There are also mechanisms, via less direct pathways through BG, for inhibiting the selection of patterns that would represent poor choices in action selection, as discussed in Gurney et al [2001] and in Houk & Wise [1995].

While a small set of potential patterns is being coarsely selected by a BG loop, a pair of loops (Figure 6) through CB amplify and sculpt this preliminary activation into a refined output vector [Houk & Mugnaini 2003]. The Nonlinear Amplification step is implemented by a positive-feedback loop through the cerebellar nucleus. Regenerative positive feedback causes Nonlinear Amplification that includes an induction of bistability in individual microscopic modules, and multistability in the mesoscopic network comprised of many microscopic modules [Wang et al 2008]. Bistability creates a behavioral threshold for the initiation of an action command while positive feedback amplifies the command's intensity, duration and spatial extent [Houk et al 1993]. Selective restraint of this amplification process and sculpting it into an accurate action command is implemented by the Refinement operation, mediated by Purkinje cells in the cerebellar cortex, considered to be the site of the brain's most powerful neuronal architecture for resolving difficult control problems [Houk & Mugnaini 2003; Houk & Wise 1995; Raymond et al 1996]. The cerebellar cortex is well suited for this important Refinement operation.

In the course of sensorimotor learning, the cerebral cortex, basal ganglia and cerebellum work in parallel but unique ways [Houk 2010; Doya 1999; Houk and Wise 1995]. The loop through the basal ganglia learns to use coarse action selection to discover ballpark actions that are appropriate in a given context, utilizing reinforcement learning [Sutton & Barto, 1998]. The loop through the cerebellum learns to refine these coarse action selections through a simplified form of supervised learning [Berthier et al 1993]. The cerebral cortex, being regulated by input from both

basal ganglia and cerebellum, learns through practice to perform these operations faster and more accurately, utilizing unsupervised Hebbian learning [Merzenich et al 1996]. All three of these adaptive mechanisms work simultaneously to implement motor learning, and their visualization using functional MR imaging (fMRI) will be a special presentation at this meeting [Scheidt et al 2011].

Pursuing this learning issue brings us back to the problem of “catastrophic interference.” Whenever a connectionist network is trained to learn something new, it necessarily forgets a little of what it knew before. This dilemma is called interference, and it can become severe if the network is overloaded. An SRN can transition to a state where it forgets just about everything it knew before. This is called catastrophic interference. Our next speaker Hiro Ohta will explain this more fully. He will claim that catastrophic interference is caused by the fixed overall representation of information and suggest that the unit of ‘distributed representation’ should be reconsidered. He will also describe how presynaptic inhibition can mediate effective long-term memory approximating incremental learning. [\[Talk 4\]](#) Presynaptic inhibition and incremental learning in the striatum of the basal ganglia, Hiro Ohta, Yasuhiro Nishida & Jim Houk.

Our next speaker Dave Fraser will introduce us to Replicate, a motor sequence task that he used to study the pathophysiology of schizophrenia. Having found a large deficit in patients, he decided to study this serial order task in other ways, including an imaging study in normal subjects. The latter results included a surprising finding, which was best explained by postulating, and modeling, pattern classification in the basal ganglia based on *presynaptic inhibition*. The title of his talk relates closely to the theme of this workshop. [\[Talk 5\]](#) Motor syntax disorder in schizophrenia, Dave Fraser & Jim Houk.

4 Special roles for presynaptic inhibition in pattern classification

Presynaptic inhibition was discussed by both of the last two speakers, but in different contexts. Dave Fraser stressed how presynaptic inhibition would improve the *performance* of the loop through the basal ganglia, whereas Hiro Ohta stressed the role of presynaptic inhibition in *learning* sequences. As an attempt to unite these two conceptual approaches, I ask the reader to think about two spiny neurons in the striatum competing with each other in order to select the next action in a sequence. If they compete by using postsynaptic inhibition, both cells will receive a large amount of synaptic input, regardless of which one wins the competition. The active synapses of the winner will be incremented and the active synapses of the loser will be decremented. In contrast, if they compete with each other by using presynaptic inhibition, one of the spiny neurons will be a “big time” winner, and it will win by turning off much of the synaptic input to the loser. The winning synapses will be incremented, but, because of the nonlinear shape of the cellular learning rule

[Shouval, Bear & Cooper 2002], the loosing synapses may not be changed at all. This example makes it easy to think through the learning that goes on at the cellular level. It makes quite clear how presynaptic inhibition could facilitate incremental learning.

5 What are motor syntax agents?

Attempting to answer this question directly, one might say “motor syntax agents are brain agents whose operations encompass motor syntax.” This takes us back to the definition of motor syntax as *the set of principles that govern sequential combinations of individual movements*. In order to motivate the notion of motor syntax, I added *in order to achieve a goal*. The goals illustrated by workshop speakers up to this point include (i) reaching to a single target rapidly and accurately (Neil Berthier), (ii) selecting the target that will reach the top of the screen next (Paul Reber), (iii) replicating the serial order of a set of K targets (Dave Fraser), (iv) satisfying thirst (Jim Houk).

Linguists usually do not consider a goal to be important to syntax. Is the existence of a goal really key to the concept of motor syntax? For language, syntax is simply structure. The obvious goal of implementing effective communication is usually not explicitly included in one’s concept of language syntax. But maybe it should be. What about during human development? The acquisition of language in the infant clearly appears to be driven by goals.

It seems that a goal is key to the definition of any given motor syntax. We also know that, when there is no external entity that specifies a goal, reinforcement learning is the mechanism the brain uses. Reinforcement learning discovers cues that predict the likelihood of future reward. These cues, together with their context, are what characterizes the “internal goal” that drives animal (including human) behavior. Since the loop through the basal ganglia is believed to mediate reinforcement learning, the sought-after motor syntax agent should include a loop through the basal ganglia. Then, in order to refine the coarse selection made by the basal ganglia loop, the DPM also needs a loop through the cerebellum.

We are coming to the conclusion that a DPM brain agent performs the role of a generic motor syntax agent. Better yet, could a DPM more generally be a *brain syntax agent*? Let’s move on to the remaining talks in this workshop, to see if they shed further light on this fundamental question.

Acknowledgement

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

Bibliography

- [1] Barringer C, Barto AG, Houk JC. Simulated reaching supports the Normalized Predicted Error (NPE) discrete control hypothesis for on-line error correction of voluntary movements. Society for Neuroscience. November 2008, Washington D.C.
- [2] Barto AG, AH Fagg, N Sitkoff and JC Houk (1999). A cerebellar model of timing and prediction in the control of reaching. *Neural Computation* 11: 565-594.
- [3] Berthier, N. E. (2011). The syntax of human infant reaching. *Proceedings of the Eighth International Conference on Complex Systems (this volume), NECSI*, in press.
- [4] Berthier N. E., S. P. Singh, A. G. Barto, and J. C. Houk (1993) Distributed representation of limb motor programs in arrays of adjustable pattern generators *J. Cognitive Neuroscience* 5, 56-78
- [5] Doya K (1999) What are the computations of the cerebellum, the basal ganglia and the cerebral cortex? *Neural Networks* 12, 961-974
- [6] Elman, J.L. An alternative view of the mental lexicon. *Trends in Cognitive Sciences* 8: 301-306, 2004.
- [7] Fishbach, A., S.A. Roy, C. Bastianen, L.E. Miller and J.C. Houk (2005). Kinematic properties of on-line error corrections in the monkey. *Exp. Br. Res.*: 164: 442-457.
- [8] Fishbach, A., S.A. Roy, C. Bastianen, L.E. Miller and J.C. Houk (2007). Deciding when and how to correct a movement: discrete submovements as a decision making process. *Exp. Br. Res.* 177: 45-63.
- [9] 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.
- [10] Gdowski MJ, Miller LE, C Bastianen, Nenonene EK, Houk JC (2007) Signaling patterns of globus pallidus internal segment neurons during forearm rotation. *Brain Research* 1155: 56-69.
- [11] Gurney, K., T. J. Prescott and P. Redgrave (2001). "A computational model of action selection in the basal ganglia. I. A new functional anatomy." *Biol. Cybern.* 84: 401-410.
- [12] Hauser M.D., Chomsky, N. and Fitch, W.T. (2002). The faculty of language: what it is, who has it, and how did it evolve. *Science* 298: 1569-1579.
- [13] Houk, J. (2001). Neurophysiology of frontal-subcortical loops. *Frontal-Subcortical Circuits in Psychiatry and Neurology*. D.G. Lichter and J. L. Cummings. New York, Guilford Publications: 92-113.
- [14] Houk JC (2005). Agents of the Mind. *Biological Cybernetics* 92: 427-437.

- [15] Houk JC (2007) Models of Basal Ganglia. Scholarpedia, p.22663. http://www.scholarpedia.org/article/Models_of_Basal_Ganglia
- [16] 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.
- [17] Houk, J. C. (2011). Action selection and refinement in subcortical loops through basal ganglia and cerebellum. Modelling Natural Action Selection. A. K. Seth, J. Bryson and T. J. Prescott, Cambridge University Press, Cambridge; in press.
- [18] Houk JC, Bastianen C, Fansler D, Fishbach A, Fraser D, Reber PJ, Roy SA, Simo LS (2007). Action selection in subcortical loops through the basal ganglia and cerebellum. Phil. Trans. R. Soc. B 362: 1573-1583
- [19] Houk, J. C., A. H. Fagg and A. G. Barto (2002). Fractional power damping model of joint motion. Progress in Motor Control: Structure - Function relations in voluntary movements. M. Latash. Champaign, Human Kinetics. 2: 147-178.
- [20] Houk, J. C., J. Keifer and A. G. Barto (1993). Distributed motor commands in the limb premotor network. Trends in Neuroscience 16: 27-33.
- [21] Houk, J and E Mugnaini (2003). Cerebellum. Fundamental Neuroscience. L.R. Squire et al., Academic Press: 841-872.
- [22] Houk, J. C. and S. P. Wise (1995). Distributed modular architectures linking basal ganglia, cerebellum, and cerebral cortex: Their role in planning and controlling action. Cerebral Cortex 5: 95-110.
- [23] Karniel, A. and G. F. Inbar (1999). The use of a nonlinear muscle model in explaining the relationship between duration, amplitude, and peak velocity of human rapid movements. J. Motor Behavior 31(3): 203-206.
- [24] Kawato, M., M. Katayama, H. Gomi and Y. Koike (1992). Coordinated arm movements: Virtual trajectory control hypothesis and learning inverse models. International Symposium on Information Sciences, Iizuka Kyusyuu, Japan.
- [25] Kelly, R. M. and P. L. Strick (2000). "Rabies as a transneuronal tracer of circuits in the central nervous system." J Neurosci Methods 103(1): 63-71.
- [26] Kelly, R. M. and P. L. Strick (2003). Cerebellar loops with motor cortex and prefrontal cortex of a nonhuman primate. J. Neuroscience 23: 8432-8444.
- [27] Kelly, R. M. and P. L. Strick (2004). Macro-architecture of basal ganglia loops with the cerebral cortex: use of rabies virus to reveal multisynaptic circuits. Prog Brain Res 143: 449-59.
- [28] Koerding, K. (2007). Decision theory: what "should" the nervous system do? Science 318(5850): 606-10.
- [29] McClelland, J. L., M. M. Botvinick, D. C. Noelle, D. C. Plaut, T. T. Rogers, M. S. Seidenberg and L. B. Smith (2010). "Letting structure emerge: connectionist and dynamical systems approaches to cognition." Trends Cogn Sci 14(8): 348-56.

- [30] Merzenich, M., B. Wright, W. Jenkins, C. Xerri, N. Byl, S. Miller and P. Tallal (1996). Cortical plasticity underlying perceptual, motor, and cognitive skill development: implications for neurorehabilitation. *Cold Spring Harb Symp Quant Biol* 61: 1-8.
- [31] Nicola, S. M., J. Surmeier and R. C. Malenka (2000). Dopaminergic modulation of neuronal excitability in the striatum and nucleus accumbens. *Annu Rev Neurosci* 23: 185-215.
- [32] Novak K.E., L.E. Miller, and J.C. Houk (2000) Kinematic properties of rapid hand movements in knob turning task. *Exp. Brain Res.* 132 (4), 419-433.
- [33] Novak, K. E., L. E. Miller and J. C. Houk (2002). The use of overlapping submovements in the control of rapid hand movements. *Exp. Brain Res.* 144: 351-364.
- [34] 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), NECS*, in press.
- [35] Patel, A. D. (2003). Language, music, syntax and the brain. *Nat Neurosci* 6(7): 674-81.
- [36] 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.
- [37] Raymond, J. L., S. G. Lisberger and M. D. Mauk (1996). The cerebellum: A neuronal learning machine? *Science* 272: 1126-1131.
- [38] Reber, P. J., D. J. Sanchez and E. W. Gobel (2011) Models of sequential learning. *Proceedings of the Eighth International Conference on Complex Systems (this volume), NECSI*, in press.
- [39] Rebuschat, P., M. Rohrmeier, J. Hawkins and I. Cross (2011). *Language and Music as Cognitive Systems*, Oxford: Oxford University Press, in press.
- [40] Redgrave, P., T. J. Prescott and K. Gurney (1999). The basal ganglia: A vertebrate solution to the selection problem? *Neuroscience* 89(4): 1009-1023.
- [41] Scheidt, R. A., N. M. G. Salowitz, J. Zimelman, A. Suminski, K. Mosier, L. S. Simo and J. C. Houk (2011) Learning visualized in cerebral cortex, basal ganglia and cerebellum. *Proceedings of the Eighth International Conference on Complex Systems (this volume), NECSI*, in press.
- [42] Shouval, H. Z., M. F. Bear and L. N. Cooper (2002). "A unified model of NMDA receptor-dependent bidirectional synaptic plasticity." *Proc Natl Acad Sci U S A* 99(16): 10831-6.
- [43] Sutton RS and AG Barto (1998) *Reinforcement Learning. An Introduction*. MIT Press, Cambridge, MA

- [44] 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.
- [45] Wang J, G Dam, S Yildirim, W Rand, U Wilensky and JC Houk (2008) Reciprocity between the cerebellum and the cerebral cortex: Nonlinear dynamics in microscopic modules for generating voluntary motor commands. *Complexity* 14 (2), 29-45.
- [46] Woodworth RS (1899) The accuracy of voluntary movement. *Psychological Review – Series of Monograph Supplements* 3 (2), 1-114.