1. Electrophysiology of abstract rule learning in 5 month-old infants

Claire Kabdebon, *Ghislaine Dehaene-Lambertz*

Cognitive Neuroimaging Unit (UNICOG) Inserm : U992, CEA-Saclay, France

The ability to learn abstract regularities from a limited set of particulars is a powerful cognitive tool that crucially comes into play in the acquisition of the rules of language. By 7 months of age, infants show surprising abilities to learn word structure in simple artificial languages, but the underlying mechanisms allowing such performances are still unclear. A growing body of research is now developing an interesting and realistic model for human learning: the Bayesian framework. In the present study, using electroencephalography, we study the brain dynamics underlying the formation of abstract representations for simple algebraic rules, and we investigate the possible neural implementation of these Bayesian mechanisms during learning. During familiarization, we presented 5 month-old infants with two types of artificial speech sequences (AAB and ABA), systematically followed by two different visual items (one for each structure). During test, infants were presented with the same audiovisual stimulation, but in 25% of the trials, we introduced an incongruency: a given structure was followed by the image associated to the other structure. From the visual evoked responses, we were able to evidence a congruency effect. This demonstrates that 5 month-olds are able to represent abstract regularities, and these representations modulate their responses to the incoming sensory inputs. We additionally investigate the neural underpinning of such abstract representations, and we try to capture the temporal dynamics of auditory evoked responses to structures using an ideal observer rule model.

2. Learning from random sequences: MEG reveals the computational fingerprints of a Bayesian inference

Maxime Maheu, *Florent Meyniel, Stanislas Dehaene*

Cognitive Neuroimaging Unit (UNICOG) Inserm : U992, CEA-Saclay, France

The brain is a highly predictive machine: it continuously computes expectations about the external world and compares them with actual events to compute prediction errors. Such errors are key components of primate cognition because they are used to revise beliefs about the world and thus underpin learning. When such prediction errors are generated, strong error signals are produced in the brain. Such “surprise” signals, among those the MMN and the P300, have been successfully elicited in experimental context as simple as listening to random binary sequence. Previous studies using this kind of design were both able to pinpoint several principles governing the elicitation of such signals as well as proposing a mathematical account explaining their amplitude as a function of stimuli history. Nevertheless, the exact properties and functions of these prediction error signals are still unknown. Here we provide a unifying theory of how the brain performs high-level inference by relying on the systematic description of these surprise signals. To do so, we first show that the surprise elicited by a particular stimulus and reflected in those error signals, depends on its predecessor. From this observation, we were able to conclude that the brain is sensitive to contextual effects when performing inference processes. We then developed Bayesian observers with different algorithmic properties and fitted their predictions to MEG signals. By describing each of the computational fingerprints we managed to isolate with this approach, we were able to show that different parts of the signal (characterized both in space and time) carry different aspects of the inference both in term of their algorithmic specifications and their substrate. Taken together these results suggest that the brain approximate Bayesian inference when facing with sequence of stimuli and that this inference relies on a hierarchy of prediction errors harboring different levels of abstraction.

3. Electrophysiology of the Bayesian brain

<u>B Kopp</u>, A Kolossa, T Fingscheidt

Univ. Braunschweig, German

Empirical support for the Bayesian brain hypothesis, although of major theoretical importance for cognitive neuroscience, is surprisingly scarce. This hypothesis posits simply that neural activities code and compute Bayesian probabilities. Here, we introduce an urn–ball paradigm to relate event-related potentials (ERPs) such as the P300 wave to Bayesian inference. Bayesian model comparison is conducted to compare various models in terms of their ability to explain trial-by-trial variation in ERP responses at different points in time and over different regions of the scalp. Specifically, we are interested in dissociating specific ERP responses in terms of Bayesian updating and predictive surprise. Bayesian updating refers to changes in probability distributions given new observations, while predictive surprise equals the surprise about observations under current probability distributions. Components of the late positive complex (P3a, P3b, Slow Wave) provide dissociable measures of Bayesian updating and predictive surprise. Specifically, the updating of beliefs about hidden states yields the best fit for the anteriorly distributed P3a, whereas the updating of predictions of observations accounts best for the posteriorly distributed Slow Wave. In addition, parietally distributed P3b responses are best fit by predictive surprise. These results indicate that the three components of the late positive complex reflect distinct neural computations. As such they are consistent with the Bayesian brain hypothesis, but these neural computations seem to be subject to nonlinear probability weighting. We integrate these findings with the free-energy principle that instantiates the Bayesian brain hypothesis.

4. Bayesian modeling of visual attention in word recognition: simulating optimal viewing position

<u>Thierry Phenix</u> , Julien Diard, Sylviane Valdois

Laboratoire de psychologie et neurocognition (LPNC) CNRS : UMR5105, Université Pierre Mendès-France - Grenoble II, Université Joseph Fourier - Grenoble I, Université de Savoie, France

In the literature, computational models of reading aloud are typically distinct from word recognition models. The former focus on high level phonological (and semantic) processes to explain how pronunciation is computed from a complete and perfectly identified sequence of letters; the latter focus on the orthographic processing stage and mainly address letter order encoding issues. None of these models includes any computational visual attention component. Based on recent behavioural and neuro-imaging evidence that visual attention span contributes to letter string processing, we develop the first model of word recognition that incorporates visual attention as a key processing device. Another originality of the model relies on the use of probabilistic programming methodologies, coming from robotics and artificial intelligence, to define modular structures in a Bayesian framework. The result is a dynamic Bayesian word recognition model, integrating bottom-up visual identification of letters and lateral interference between visual inputs with top-down control of attentional resources and lexical knowledge. (We call the model BRAID, for Bayesian word Recognition with Attention, Interference and Dynamics.) In a series of stimulations, we first show how visual attention modulates letter identification and letter position encoding. We will report simulations of classical effects of frequency, word superiority, priming or letter transposition. BRAID further allows simulating more challenging behavioural data such as the optimal viewing position effect, due to the possibility to focus attention on different letters of the input word.

5. Modeling the concurrent development of speech perception and production in a Bayesian framework

Marie-Lou Barnaud, Julien Diard, Pierre Bessière, Jean-Luc Schwartz

Grenoble Images Parole Signal Automatique (GIPSA-lab) Université Stendhal - Grenoble III, Université Pierre Mendès-France - Grenoble II, Université Joseph Fourier - Grenoble I, CNRS : UMR5216, Institut Polytechnique de Grenoble, France

It is widely accepted that both motor and auditory processes interact in the brain during speech perception, but little is known about the functional role played by motor processes. To address this question we consider a Bayesian model of speech communication based on three sets of variables: motor representations M, sensory representations S and objects O (e.g. phonological units such as phonemes). The model comprises two internal branches. Firstly, an auditory identification sub-system connects S and O. Secondly, a motor production sub-system connecting M and O and a sensory-motor sub-system connecting M and S can be combined to provide "motor identification" of sounds S, from S to M and from M to O, in an analysis-by-synthesis process.

The auditory identification sub-system, the motor production sub-system and the sensory-motor sub-system are learned in a supervised learning scenario, in which a master agent provides sensory signals s and their respective object o. Learning the auditory identification system is straightforward using experimental < s; o > pairs. On the other hand, learning the motor sub-system is more complicated. The learning agent infers motor gestures in an "accomodation process": the learning agent tries to reproduce the input sensory signal s by selecting a motor gesture m. Performing m yields s', the resulting sensory output. Triplets < m; s'; o > are used to update the parameters of the motor identification system.

We show that the direct inference process involved in auditory identification provides rapid and efficient learning but generalizes poorly. By contrast, the more complex inference process required in motor identification learns more slowly and performs less accurately. However, this system happens to have captured more variable situations during learning, and generalizes better (e.g. in noise). This could provide the basis for a complementarity between auditory and motor identification systems in the human brain.

6. A Bayesian framework for speech motor control

Jean-François Patri, Julien Diard, Jean-Luc Schwartz [1], *Pascal Perrier* [2]

1 : Grenoble Images Parole Signal Automatique (GIPSA-lab) Université Stendhal - Grenoble III, Université Pierre Mendès-France - Grenoble II, Université Joseph Fourier - Grenoble I, CNRS : UMR5216, Institut Polytechnique de Grenoble, France

2 : Grenoble Images Parole Signal Automatique (Gipsa-lab) Grenoble Images Parole Signal Automatique

Speech is a skilled serial-order motor task achieving time series of goals within a timing that does not allow any online cortical processing of feedback signals. In addition, the speech motor system is highly redundant with many available degrees of freedom, which makes the inference of motor commands from the physical signals an "ill-posed" inverse problem.

To deal with this complexity, speech planning has been classically modeled within an optimal motor control framework considering a feedforward control scheme coupled with a feedback controller enabling a correction of motor commands. This approach has proven to generate results in close agreement with experimental data, in particular in terms of adaptation to perturbations or in terms of anticipatory behavior.

However, criticisms of this approach concern in particular the inability to account for the well-known token-to-token speech variability. This drawback is inherent to the feedforward optimal control scheme, since it basically cancels all possible variations along the degrees of freedom directions, by specifying a unique optimal solution to the control problem.

In the present work, we propose an alternative approach by formulating feedforward optimal control in a Bayesian modeling framework. We consider this approach to be appropriate for solving the ill-posed inverse problem while accounting for the observed token-to-token variability in a principled way. The basic principles underlying the search for optimality is preserved without being explicitly driven by the minimization of a cost.

The approach is illustrated by reformulating an existing optimal control model for speech planning into the Bayesian modeling framework. We demonstrate that models are nested, with optimal control as a special case of the Bayesian model. Variability is formally generated by assuming that control is performed by sampling motor commands randomly according to the distribution solving the inference problem.

7. Biochemical probabilistic computation

David Colliaux, *Jacques Droulez and Pierre Bessière*

ISIR, UPMC Paris, France

To face an ever changing and uncertain environment and to behave adaptively, living organisms rely on probabilistic strategies. This has led to the hypothesis that Bayesian computations are embedded in the brain. Most of the proposed models used neurons as the atomic elements composed in assemblies to perform more complex computations. In complementation of these models at the neuronal scale, we propose that probabilistic inference could also take place at the molecular scale, using biochemical signalling as the backbone for the required computations. On one hand, we show that any Bayesian inference problem on discrete finite variables can be mapped to the computation of a rational function with non-negative coefficient (RFNC). We also show the reciprocal of this proposition. On the other hand, by studying models of feed forward biochemical cascades, we show that the stationary states of products are RFNNC of the input messenger concentrations. Reciprocally, we demonstrate that for any RFNC there exist a theoretical biochemical cascade that may computes it. Finally, these 4 propositions put together proves that cell signalling can be considered as a possible alternative mechanism for probabilistic inference. It obviously augments the available computing power for organisms with a nervous system and extends the possibility to treat uncertain information to unicellular organism. This hypothesis will be illustrated by simulations of biochemical cascades performing some basic Bayesian computation.

8. The emergence of emotional valence under active inference

Mateus Joffily

Groupe d'analyse et de théorie économique (GATE Lyon Saint-Étienne) École Normale Supérieure - Lyon, Université Lumière - Lyon II, Université Claude Bernard - Lyon I (UCBL), CNRS : UMR5824, Université Jean Monnet - Saint-Etienne, PRES Université de Lyon, France

Previous work has suggested that emotional valence – the positive or negative attribute of affect, emotions and their associated feelings – could be formally defined as the rate of change of free energy or surprise (Joffily and Coricelli, 2013). An apparent paradox emerged regarding how ‘positive surprises' and ‘negative expectations' can be accounted for within this scheme. The aim of the current work is to show that these two forms of affective states naturally emerge when the agent performs active inference. Active inference rests on the active sampling of outcomes that minimize variational free energy (Friston et al, 2015). Based on simulation results for a two-armed bandit task, it is shown that the valence of outcomes can be described as the difference between the current and the expected free energy under a (predictive) generative model of future states and outcomes, while the surprise of outcomes is formulated as the free energy of posterior beliefs about their hidden causes.

9. Integration of beliefs and affective values in human decision-making

Marion Rouaul [1], *Jan Drugowitsch* [2], *Anne-Dominique Lodeho* [1], *Etienne Koechlin* [1]

1 : Laboratoire de Neurosciences Cognitives INSERM U960 Ecole Normale Supérieure de Paris, France

2 : Departement des Neurosciences Fondamentales, Universite de Geneve , Switzerland

Executive control relies on evaluating action outcomes for adjusting subsequent action. Action outcomes, however, may convey two types of value signals:

- Affective values (Av), representing the valuation of action outcomes given subjective preferences
- Belief values (Bv), about how actions map onto outcome contingencies.

Av stems from reinforcement learning (RL), whereas Bv stems from Bayesian models. Previous experimental paradigms usually confounded Av and Bv: a higher reward usually informs about more appropriate choices.

We present a probabilistic reversal adaptive task aiming at dissociating Av from Bv. Human subjects had to decide between two shapes, one of which was more frequently rewarded than the other one. The potential rewards to gain for each shape were displayed before each choice. Crucially, we manipulated the reward distributions underlying each shape to dissociate Av from Bv. We developed a computational model of monitoring and decision, integrating two parallel systems: RL, dealing with Av, and Bayesian inference, dealing with Bv. This model fitted behavioral performance better than many alternative models.

We then investigated whether beliefs and affective values had distinct neural bases using fMRI. BOLD signal was regressed against choice-dependent and choice-independent beliefs and affective values. Ventromedial prefrontal cortex (VMPFC) and anterior cingulate cortex (ACC) activity correlated with both choice-dependent variables. However, we found a double-dissociation regarding choice-independent variables, with VMPFC encoding choice-independent beliefs, whereas ACC encoded choice-independent affective values. Additionally, activity in lateral prefrontal cortex (LPFC) increased when decision values (i.e. mixture of beliefs and affective values) got closer to each other and action selection became more difficult.

Taken together, these results suggest that before decision, VMPFC and ACC separately encode beliefs and affective values respectively. LPFC combines both signals to decide, then feeds back choice information to the medial regions, presumably for updating these value signals according to action outcomes.

10. How do beliefs affect percepts in schizophrenia? An experimental validation of the circular inference framework

Leptourgos P., Notredame C. E., Jardri R., Denève S.

(ENS, Paris, France)

Computational psychiatry is here! During the last years, more and more scientists approach psychiatric disorders by focusing their efforts on explaining the computational mechanisms that might cause them. Especially the positive dimension of schizophrenia (i.e. hallucinations and delusional ideations) has been interpreted as an impairment in how sensory information is integrated with prior beliefs, a deficit that can be formalized as impaired Bayesian inference. We propose that an imbalance between excitation and inhibition in brain networks (cortical and subcortical) leads to circular inference, i.e. an aberrant form of inference where messages (bottom up and / or top down) are counted more than once and thus, are overweighted (Jardri and Deneve (2013)). More specifically, we postulate that psychotic symptoms are caused by a system that “expect what it senses” and as a result attributes extreme weight even to weak sensory evidences.

A previous experiment validated our hypothesis in a probabilistic inference task. Interestingly, we found significant amounts of circular inference in healthy controls as well, albeit to a lesser extent. Here, we put forward a new experimental study that could validate the circular inference framework in the domain of visual perception. As a first step, we restricted ourselves to healthy controls, whose tendencies for psychotic symptoms were measured using appropriate scales. We investigated the computations performed by perceptual systems when facing ambiguous sensory evidence (Necker Cube). In cases of high ambiguity, perception is unable to create stable interpretations of the input and oscillates between two mutually exclusive unstable interpretations, a phenomenon known as bistable perception. More specifically, we asked how prior expectations and visual cues can affect the dynamics of bistability. Participants were asked to look at a 2D Necker cube that was continuously displayed on the screen and discontinuously report their percept (3D cube) every time they hear a sound. We manipulated sensory evidence by adding shades to the stimuli and prior expectations by giving different instructions to different groups of participants, concerning the presence of an implicit bias. We showed that both prior expectations and visual cues (and their interactions) significantly affect the pattern of bistability, using both static (alternation rate, relative predominance, phase durations) and dynamic measures (transient preference, reversal probability, survival probabilities). We found that the behavior could be well fitted by Bayesian models (“simple” Bayes, hierarchical Bayesian model with Markovian statistics) with low statistical dependencies between successive time steps. The above findings could be used as a reference, in order to extend the study to include patients with psychotic symptoms and ultimately test our initial claim that psychosis is the result of over-counted sensory evidence.

11. Dissociating sensory from decision processes in human perceptual decision making

Pim Mostert, Peter Kok, Floris P. de Lange

Radboud University, Donders Institute for Brain, Cognition and Behaviour, 6500 HB, Nijmegen, The Netherlands

Theoretical frameworks posit that perceptual decisions arise from a cascade of functionally distinct stages. In the sensory stage, the physical stimulus is encoded into internal sensory evidence, while during the subsequent decision stage, this sensory evidence is integrated over time into a decision variable. Identifying and dissociating these stages at the neural level will advance our understanding of perceptual decision making and has become a central goal within the field.

We adopted a novel approach to dissociate sensory- from decision-related activity in MEG, while subjects were making perceptual decisions about the presence/absence of a grating embedded in noise. We used a functional localizer, in which subjects were exposed to grating stimuli while performing a task at fixation, in order to isolate neural signals related to sensory processing in the absence of decision making and attention. A multivariate decoder was trained on this data to identify a sensory-specific neural signature. We subsequently employed this decoder to trace sensory processing over time during perceptual decision making. In addition, we conducted an analysis in which we trained a decoder on the data obtained during the perceptual decision making task itself, without making use of the functional localizer, to extract the neural signal that collectively underlies sensory processing and decision making.

The results showed that sensory-related activity was specific to early time windows and consistent with occipital sources. Interestingly, we observed that the stimulus information was sustained when it was attended and relevant for the task at hand. Furthermore, the stimulus information faithfully reflected the physical stimulus, regardless of the eventual behavioral decision. Decision-related activity, on the other hand, was longer-lasting and more prominent during later time windows. It showed a ramping temporal profile and was localized to parietal and frontal cortex. Furthermore, decision-related activity was confined to a later time window when no grating was presented, but included the early time points when a grating was presented - the time window that was previously associated with sensory processing. We suggest that this early decision-related activity may reflect fluctuations in latent confounding variables, such as attention.

In conclusion, our approach provides a novel way to reliably extract the neural dynamics of sensory processing during perceptual decision making, uncontaminated by decision processes or other confounding variables.

12. A Bayesian model of pain self-report shows variability in the influence of prior expectations between individuals

Hoskin, R., Berzuini, C., Guo, H. & Talmi, D.

University of Manchester, UK

When we expect pain to be intense we experience it as stronger than when we expect it to be less intense. An established research literature has shown that when people receive the same physical intensity of pain stimulation they report a higher, or lower, pain experience when they expect a higher or lower pain, respectively. We do not know, however, exactly what people care about or pay attention to when evaluating future pain. This is an obstacle when translating existing research to clinical settings. Here we describe a new experimental task where we varied the mean and variance of expected pain in healthy volunteers, and then delivered pain in the form of electric stimulation of the skin of the hand. We examined how volunteers' pain ratings was affected by prior expectation, time and delivered stimulus through a Bayesian hierarchical model that allows for individual-level heterogeneity. The random effects structure of the model allowed individual-specific estimates of the effect of expectation to be distilled from the data. These estimates we regard as individual-level measures of sensitivity to expectancy, and call Individual Expectation Scores (IESs). The proposed model, in combination with a Markov Chain Monte Carlo (MCMC) approach to inference, produced Bayesian credible intervals for the IESs, as a measure of their uncertainty. Expectation exerted on average a statistically significant effect on the rating, over and above that of delivered pain level. Important differences between the IESs of different individuals emerged from the analysis.

13. Optimal visuo-tactile integration for velocity discrimination of self-hand movements

Chancel M. [1,2]*, Blanchard C.* [3]*, Guerraz M.* [2]*, Montagnini A.* [4]*, Kavounoudias A.* [1]

1. Aix Marseille Université, CNRS, NIA UMR 7260, 13331 Marseille, France

2. Université Savoie Mont Blanc, CNRS, LPNC UMR 5105, F-73000 Chambéry, France

3. University of Nottingham, School of Psychology, NG7 2RD Nottingham, UK

4. Aix Marseille Université, CNRS, INT UMR 7289, 13385 Marseille, France

We recently showed that illusory hand movements can be elicited by either a textured disk or a visual pattern rotating under one's hand while proprioceptive inputs convey immobility information (Blanchard et al., 2013). Here we investigated how visuo-tactile integration can optimize velocity discrimination of illusory hand movements and whether the Bayesian Maximum Likelihood Estimation (MLE) model can predict such an effect. We induced illusory movements in eleven volunteers by visual and/or tactile stimulation, randomly delivered at 6 angular velocities. The participants then had to compare hand illusion velocities with a 5°/s hand reference movement in an alternative forced choice paradigm. Results showed that the discrimination threshold decreased under visuo-tactile condition compared to unimodal (visual or tactile) conditions, reflecting better bimodal discrimination. The gains of the illusions also increased: the stimulation required to give rise to a 5°/s illusory movement was slower in the visuo-tactile compared to each of the two unimodal conditions. The MLE satisfactorily predicted the improved discrimination threshold, but not the increase in gain. When we added a zero-centered Prior, reflecting immobility information, the Bayesian model did actually predict (though over-estimated) the gain increase. Interestingly, the predicted gains better fit the visuo-tactile performances when a proprioceptive noise was generated by co-vibrating antagonist wrist muscles. These findings suggest that kinesthetic information of visual and tactile origins is optimally integrated to improve velocity discrimination of self-hand movements. They further highlight the relative importance of the omnipresent muscle proprioceptive cues with respect to other sensory cues for kinesthesia.

14. Active Inference, tracking eye movements and oculomotor delays

Laurent U. Perrinet [1,2]*, Rick A. Adams* [2]*, Karl J. Friston* [2]

1. Institut de Neurosciences de la Timone, CNRS/Aix-Marseille Université, France

2. The Wellcome Trust Centre for Neuroimaging, University College London, UK

Tracking eye movements face a difficult task: they have to be fast while they suffer inevitable delays. If we focus on area MT of humans for instance as it is crucial for detecting the motion of visual objects, sensory information coming to this area is already lagging some 35 milliseconds behind operational time – that is, it reflects some past information. Still the fastest action that may be done there is only able to reach the effector muscles of the eyes some 40 milliseconds later – that is, in the future. The tracking eye movement system is however able to respond swiftly and even to anticipate repetitive movements (e.g. Barnes et al, 2000 – refs in manuscript). In that case, it means that information in a cortical area is both predicted from the past sensory information but also anticipated to give an optimal response in the future. Even if numerous models have been described to model different mechanisms to account for delays, no theoretical approach has tackled the whole problem explicitly. In several areas of vision research, authors have proposed models at different levels of abstractions from biomechanical models, to neurobiological implementations (e.g. Robinson, 1986) or Bayesian models. This study is both novel and important because – using a neurobiologically plausible hierarchical Bayesian model – it demonstrates that using generalized coordinates to finesse the prediction of a target's motion, the model can reproduce characteristic properties of tracking eye movements in the presence of delays. Crucially, the different refinements to the model that we propose – pursuit initiation, smooth pursuit eye movements, and anticipatory response – are consistent with the different types of tracking eye movements that may be observed experimentally.

15. Rules and Goals: Neural mechanisms of cognitive control in human prefrontal cortex

Sandrine Duverne and Etienne Koechlin

Ecole Normale Supérieure and INSERM

The human prefrontal cortex (PFC) subserves cognitive control, *i.e.* the ability to select behavioral strategies in relation to current mental states. Cognitive control is usually postulated as being goal-directed: strategy selection follows current expectations about action outcomes based on external cues. However, recent computational proposals challenge this goal-centered view and instead describe cognitive control as inferring from both external cues and action outcomes which strategy applies to the current situation. This inference-centered view has yet to be empirically investigated. Using neuroimaging, we show here that the medial PFC encodes and conveys to lateral PFC reward expectations driving strategy selection in lateral PFC. Critically, however, strategy selection in lateral PFC could comply with external cues in contradiction to current reward expectations that medial PFC encoded and conveyed to lateral PFC. This functional coupling between medial and lateral PFC reflects inference-driven rather than goal-directed processes in the service of cognitive control.