



Special issue: Research report

Active inference, communication and hermeneutics[☆]

Karl J. Friston^{*} and Christopher D. Frith

The Wellcome Trust Centre for Neuroimaging, Institute of Neurology, UCL, United Kingdom

ARTICLE INFO

Article history:

Received 28 September 2014
 Reviewed 28 November 2014
 Revised 6 December 2014
 Accepted 27 March 2015
 Published online 15 April 2015

Keywords:

Communication
 Neuronal
 Hermeneutics
 Theory of mind
 Active inference
 Predictive coding
 Bayesian
 Synchronisation of chaos

ABSTRACT

Hermeneutics refers to interpretation and translation of text (typically ancient scriptures) but also applies to verbal and non-verbal communication. In a psychological setting it nicely frames the problem of inferring the intended content of a communication. In this paper, we offer a solution to the problem of *neural hermeneutics* based upon *active inference*. In active inference, action fulfils predictions about how we will behave (e.g., predicting we will speak). Crucially, these predictions can be used to predict both self and others – during speaking and listening respectively. Active inference mandates the suppression of prediction errors by updating an internal model that generates predictions – both at fast timescales (through *perceptual inference*) and slower timescales (through *perceptual learning*). If two agents adopt the same model, then – in principle – they can predict each other and minimise their mutual prediction errors. Heuristically, this ensures they are singing from the same hymn sheet. This paper builds upon recent work on active inference and communication to illustrate perceptual learning using simulated birdsongs. Our focus here is the neural hermeneutics implicit in learning, where communication facilitates long-term changes in generative models that are trying to predict each other. In other words, communication induces perceptual learning and enables others to (literally) change our minds and *vice versa*.

© 2015 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The term hermeneutics refers to the art of interpreting written texts such as holy scriptures. The key problem for

hermeneutics rests on developing criteria for deciding when an interpretation is correct. This problem is not restricted to the interpretation of ancient texts. When talking to you, I cannot access your mind to check whether my interpretation

[☆] “Much of the introductory material (and text) in the current paper is taken from a companion paper (Friston and Frith, 2015) that establishes the basic model of birdsong used in this paper. The companion paper establishes the notion of generalised synchrony in the setting of communication, focusing on sensory attenuation and its role in turn taking (assuming a shared generative model). The current paper addresses the next fundamental issue; namely the (joint) acquisition or learning of generative models that underlie communication and dynamic coordination (Kelso 2012).” Kelso JA. (2012). Multistability and metastability: understanding dynamic coordination in the brain. *Philos Trans R Soc Lond B Biol Sci.* Apr 5; 367(1591):906–18.

^{*} Corresponding author. The Wellcome Trust Centre for Neuroimaging, Institute of Neurology, UCL WC1N 3BG, United Kingdom.

E-mail address: k.friston@ucl.ac.uk (K.J. Friston).

<http://dx.doi.org/10.1016/j.cortex.2015.03.025>

0010-9452/© 2015 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

of what you have just said corresponds to what you meant. I can invent a coherent story or narrative, but I can never independently verify my interpretations (Frith & Wentzer, 2013). Nevertheless, people seem to understand each other most the time. How is this achieved? In this paper, we suggest the criteria for evaluating and updating my interpretation of your behaviour are exactly the same criteria that underlie action and perception in general; namely, the minimisation of prediction error or (variational) free energy.

In a companion paper (Friston and Frith, 2015), we considered communication in terms of inference about others, based on the notion that we model and predict our sensations – sensations that are generated by other agents like ourselves. This leads to a view of communication based on a generative model or narrative that is shared by agents who exchange sensory signals. Given a shared narrative, communication can then be cast as *turn taking* (Wilson & Wilson, 2005), by selectively attending and attenuating sensory information. Attending to exteroceptive sensations enables the shared narrative to predict the sensory input generated by another (while listening). Conversely, attenuating exteroceptive input enables one to articulate the narrative by realising proprioceptive predictions (while speaking). Using simulations, we demonstrated this turn taking by assuming that both agents possessed the same generative model. In this paper, we consider how and why generative models learned by agents – who exchange sensory signals – become the same (shared) model.

Our underlying premise is that we are trying to model the causes of our sensations – and adjust those models to maximise Bayesian model evidence or, equivalently, minimise surprise (Brown & Brün, 2012; Kilner, Friston, & Frith, 2007). This perspective on action and perception has broad explanatory power in several areas of cognitive neuroscience – and enjoys support from several lines of neuroanatomical and neurophysiological evidence (Egner & Summerfield, 2013; Rao & Ballard, 1999; Srinivasan, Laughlin, & Dubs, 1982). In communication and the interpretation of intent, the very notion of *theory of mind* speaks directly to inference, in the sense that theories make predictions that have to be tested against (sensory) data. Imagine two brains, each mandated to model the (external) states of the world causing sensory input. Now imagine that sensations can only be caused by (the action of) one brain on the other. This means that the first brain has to model the second. However, the second brain is modelling the first, which means the first brain must have a model of the second brain, which includes a model of the first – and so on *ad infinitum*. At first glance, the implicit infinite regress appears to preclude a veridical modelling of another's brain. However, this infinite regress dissolves if each brain models the sensations caused by itself and the other as being generated in the same way. In other words, if there is a shared narrative or dynamic that both brains subscribe to, then they can predict each other exactly – at least for short periods of time. This is the basic idea that we pursue in the context of active inference and predictive coding.

In our previous paper, we focused on the dynamical phenomena that emerge when two dynamical systems try to predict each other. Mathematically, this dynamical coupling is

called *generalised synchrony* (aka synchronisation of chaos) (Barreto, Josic, Morales, Sander, & So, 2003; Hunt, Ott, & Yorke, 1997). Generalised synchrony was famously observed by Huygens in his studies of pendulum clocks – that synchronized themselves through the imperceptible motion of beams from which they were suspended (Huygens, 1673). This nicely illustrates the *action at a distance* among coupled dynamical systems. Put simply, generalised synchronisation means that knowing the state of one system (e.g., neuronal activity in the brain) means one can predict the another system (e.g., another's brain).

We will consider a special case of generalized synchronization; namely, *identical synchronization*, in which there is a one-to-one relationship between the states of two systems. Identical synchronisation emerges when the systems that are coupled are the same. In the context of active inference, this means the two generative models are identical. But why should two agents have the same generative model? The answer is rather obvious – when they share the same generative model they can predict each other more accurately and minimise their prediction errors or surprise. The key point here is that the same principle that leads to generalised synchrony also applies to the selection or learning of the model generating predictions. This learning is the focus of the current paper, which provides an illustrative proof of principle that the hermeneutic cycle can be closed by simply updating generative models and their predictions to minimise prediction errors. Crucially, these prediction errors can be computed without ever knowing the true state of another; thereby solving the problem of hermeneutics (see Fig. 1).

The treatment of communication in this paper is rather abstract and borrows mathematical concepts from dynamical systems theory. Although we will use birdsong as a vehicle to illustrate the ideas, we do not pretend this is a meaningful model of linguistic communication (or indeed songbirds). Rather, we try to understand the dynamic coordination of richly structured behaviours, such as singing and dancing, without ascribing any (semantic) meaning or syntax to sensory exchanges. Having said this, there is growing interest in applying the principles of predictive coding to language: e.g., (Arnal & Giraud, 2012; Hickok, 2013; Pickering & Clark, 2014; Wang, Mathalon, et al., 2014) – and understanding the algebra of dynamical systems in terms of communication; e.g., (Scott-Phillips & Blythe, 2013). Furthermore, predictive coding is starting to shed light on spectral asymmetries – in coupling within the auditory hierarchy – evident in electrophysiological studies of speech processing (Arnal, Wyart, & Giraud, 2011).

This paper comprises five sections. The first sections reprise the material in (Friston and Frith, 2015), which provides a brief review of active inference and predictive coding in communication. In the second section, we described the particular (birdsong) model used to illustrate communicative inference. This model has been used previously to illustrate several phenomena in perception; such as perceptual learning, repetition suppression, and the recognition of stimulus streams with deep hierarchical structure (Friston & Kiebel, 2009; Kiebel, Daunizeau, & Friston, 2008). In the third section, we provide a simple illustration of omission related responses – that are ubiquitous in neurophysiology and

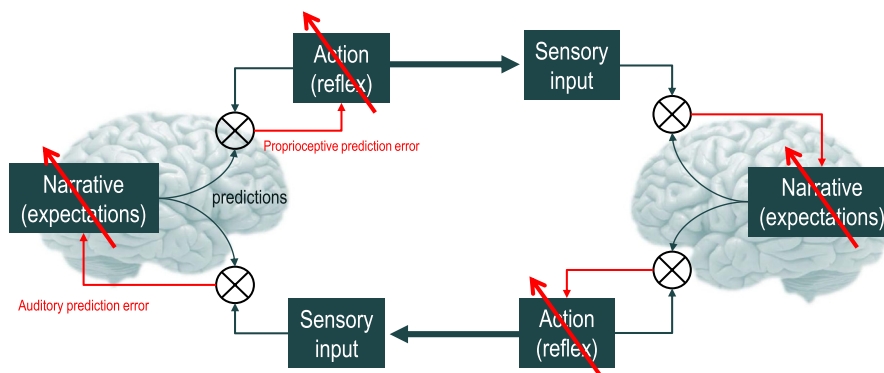


Fig. 1 – A predictive coding formulation of the Hermeneutic Circle: e.g., Speech Chain in the setting of language (Denes & Pinson, 1993). This schematic provides a simple example of neural hermeneutics in the form of a control system diagram. It is a simple example because the internal (generative) model predicting the behavioural consequences of action (of self and other) is the same. In other words, it is neither a model of my behaviour or your behaviour – but a model of our behaviour. When both agents adopt the same model, generalised synchronisation is guaranteed and prediction errors are minimised. The implicit architecture highlights the fact that the top-down predictions from a dynamical generative model (labelled Narrative) come in two flavours: exteroceptive predictions predicting the external consequences of action (c.f., corollary discharge) and proprioceptive predictions that predict the internal consequences of action (c.f., motor commands). These predictions are compared with sensory input to provide prediction errors. In control diagrams of this sort \otimes denotes a comparator. Exteroceptive (e.g., auditory) prediction errors are used to update the generative model at various timescales to produce inference and learning. In contrast, the proprioceptive prediction errors drive classical reflexes to produce the predicted action. When the (dynamics of the) generative models in the two brains are identical, both exteroceptive and proprioceptive prediction errors are minimised and the dynamics will exhibit (generalised) synchrony. The red arrows denote learning or control by prediction errors that compare (descending) predictions with (ascending) sensations.

illustrate the basic nature of predictive coding. In the fourth section, we use this model to simulate two (identical) birds that are singing to themselves (and each other) and examine the conditions under which generalised synchrony emerges. In the final section, we repeat the simulations using two birds that start off with different generative models. We will see that perceptual learning produces a convergence of the two generative models over time – leading to the emergence of generalised synchrony and implicit communication. We offer this as a solution to the problem of hermeneutic inference that can be resolved by neuronally plausible schemes – with the single imperative to minimise prediction error or free energy.

2. Active inference and predictive coding

Recent advances in theoretical neuroscience have produced a paradigm shift in cognitive neuroscience. This shift is away from the brain as a passive filter of sensations – or an elaborate stimulus-response link – towards a view of the brain as an organ that generates hypotheses or fantasies (fantastic: from Greek *phantastikos*, the ability to create mental images, from *phantazesthai*), which are tested against sensory evidence (Gregory, 1968). This perspective dates back to the notion of unconscious inference (Helmholtz, 1866/1962) and has been formalised to cover deep or hierarchical Bayesian inference – about the causes of our sensations – and how these inferences induce beliefs, movement and behaviour (Clark,

2013; Dayan, Hinton, & Neal, 1995; Friston, Kilner, & Harrison, 2006; Hohwy, 2013; Lee & Mumford, 2003).

2.1. Predictive coding and the Bayesian brain

Modern formulations of the Bayesian brain – such as predictive coding – are now among the most popular explanations for neuronal message passing (Clark, 2013; Friston, 2008; Rao & Ballard, 1999; Srinivasan et al., 1982). Predictive coding is a biologically plausible process theory for which there is a considerable amount of anatomical and physiological evidence. In these schemes, neuronal representations – in higher levels of cortical hierarchies – generate predictions of representations in lower levels (Friston, 2008; Mumford, 1992; Rao & Ballard, 1999). These top-down predictions are compared with representations at the lower level to form a prediction error (usually associated with the activity of superficial pyramidal cells). The resulting mismatch signal is passed back up the hierarchy to update higher representations (associated with the activity of deep pyramidal cells). This recursive exchange of signals suppresses prediction error at each and every level to provide a hierarchical explanation for sensory inputs that enter at the lowest (sensory) level. In computational terms, neuronal activity encodes beliefs or probability distributions over states in the world that cause sensations (e.g., my visual sensations are caused by a *face*). The simplest encoding corresponds to representing the belief with the expected value or *expectation* of a (hidden) cause.

These causes are referred to as *hidden* because they have to be inferred from their sensory consequences.

In summary, predictive coding represents a biologically plausible scheme for updating beliefs about states of the world using sensory samples: see Fig. 2. In this setting, cortical hierarchies are a neuroanatomical embodiment of how sensory signals are generated; for example, a face generates surfaces that generate textures and edges and so on, down to retinal input. This form of hierarchical inference explains a large number of anatomical and physiological facts as reviewed elsewhere (Adams, Shipp, & Friston, 2013; Bastos et al., 2012; Friston, 2008). In brief, it explains the hierarchical nature of cortical connections; the prevalence of backward connections and many of the functional and structural

asymmetries in the extrinsic connections that link hierarchical levels (Zeki & Shipp, 1988). These asymmetries include the laminar specificity of forward and backward connections, the prevalence of nonlinear or modulatory backward connections and their spectral characteristics – with fast (e.g., gamma) activity predominating in forward connections and slower (e.g., beta) frequencies that accumulate evidence (prediction errors) ascending from lower levels.

2.2. Precision engineered message passing

One can regard ascending prediction errors as broadcasting ‘newsworthy’ information that has yet to be explained by descending predictions. However, the brain has to select the

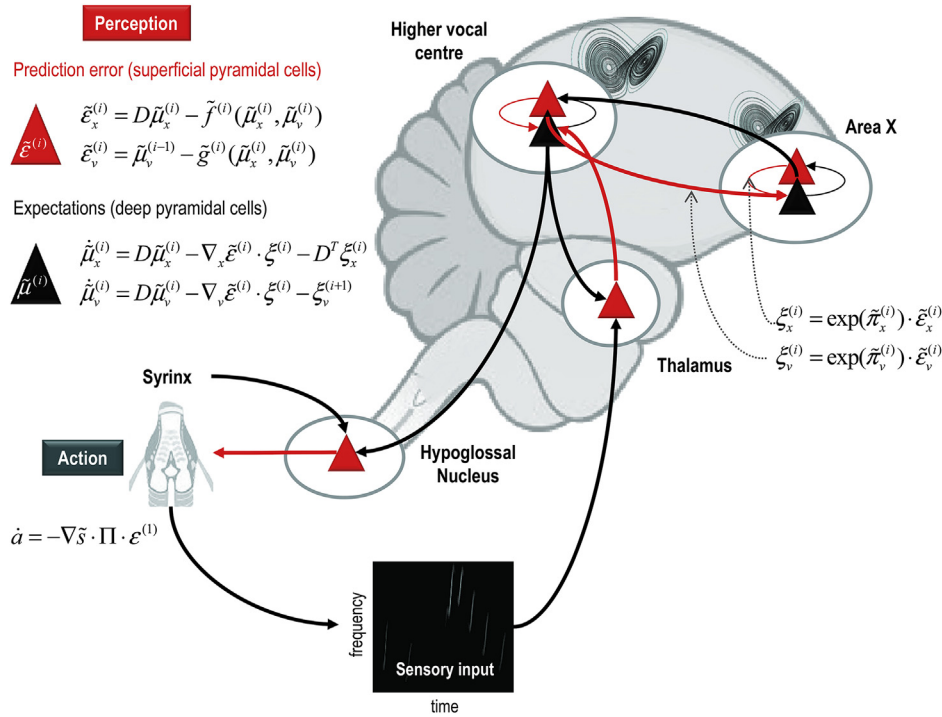


Fig. 2 – This figure summarizes hierarchical neuronal message passing in predictive coding using the (simplified) neuroanatomy of a songbird. Neuronal activity encodes expectations about the causes of sensory input, where these expectations minimize prediction error. Prediction error is the difference between (ascending) sensory input and (descending) predictions of that input. Here, sensory input is represented by a sonogram encoding the amplitude of different acoustic frequencies over time. Minimising prediction error rests upon recurrent neuronal interactions among different levels of the cortical hierarchy. The available evidence suggests that superficial pyramidal cells (red triangles) compare the expectations (at each level) with top-down predictions from deep pyramidal cells (black triangles) of higher levels. Left panel: these equations represent the neuronal dynamics implicit in predictive coding. Prediction errors at the *i*-th level of the hierarchy are simply the difference between the expectations encoded at that level and top-down predictions of those expectations. The expectations *per se* are driven by prediction errors so that they reduce the sum of squared (precision weighted) prediction error. See the appendix for a detailed explanation of these equations and the variables in this figure. Right panel: this provides a schematic example in the auditory system of a songbird: it shows the putative cells of origin of ascending or forward connections that convey (precision weighted) prediction errors (red arrows) and descending or backward connections (black arrows) that construct predictions. In this example, area X sends predictions to HVC (c.f., high vocal centre), which projects to the auditory thalamus. However, the HVC also sends proprioceptive predictions to the hypoglossal nucleus, which are passed to the syrinx to generate vocalisation through classical reflexes. These predictions can be regarded as motor commands, while the descending predictions of auditory input correspond to corollary discharge. Note that every top-down prediction is reciprocated with a bottom-up prediction error to ensure predictions are constrained by sensory information. The neuroanatomy implicit in this schematic should not be taken too seriously: we have simply transcribed a generic hierarchical message passing scheme onto the key connections among the cardinal regions implicated in the processing of birdsongs (Nottebohm, 2005).

channels it listens to – by adjusting the volume or *gain* of prediction errors that compete to update expectations in higher levels. Computationally, this gain corresponds to the precision or confidence associated with ascending prediction errors. However, to select prediction errors the brain has to estimate and encode their precision (i.e., inverse variance). Having done this, prediction errors can then be weighted by their precision so that only precise information is assimilated at high or deep hierarchical levels. The dynamic and context-sensitive control of precision has been associated with attentional gain control in sensory processing (Feldman & Friston, 2010; Jiang, Summerfield, & Egner, 2013) and has been discussed in terms of affordance in active inference and action selection (Cisek, 2007; Frank, Scheres, & Sherman, 2007; Friston et al., 2012). Crucially, the delicate balance of precision at different hierarchical levels has a profound effect on veridical inference – and may also offer a formal understanding of false inference in psychopathology (Adams, Stephan, Brown, Frith & Friston, 2013; Fletcher & Frith, 2009).

In (Friston and Frith, 2015) we illustrated the role of precision in mediating sensory attenuation and its necessary role in enabling (open loop) motor control during action. See also (Brown, Adams, Parees, Edwards, & Friston, 2013). In brief, this leads to an alternation between sensory attention and attenuation in the action–perception cycle, which involves a temporary suspension of attention to the consequences of acting during the act itself: see (Numminen, Salmelin, & Hari, 1999). We will use the same approach in this paper to model turn taking (Wilson & Wilson, 2005) in the switching between speaking (singing) and listening.

2.3. Active inference

Hitherto, we have only considered the role of predictive coding in perception through minimising surprise or prediction errors. However, there is another way to minimise prediction errors; namely, by re-sampling sensory inputs so that they conform to predictions: in other words, changing sensory inputs by changing the world through action. This is known as active inference (Friston, Mattout, & Kilner, 2011). In active inference, action is regarded as the fulfilment of descending proprioceptive predictions by classical reflex arcs. In other words, we believe that we will execute a goal-directed movement and this belief is unpacked hierarchically to provide proprioceptive, and exteroceptive predictions generated from our generative or forward model. These predictions are then fulfilled automatically by minimizing proprioceptive prediction errors at the level of the spinal cord and cranial nerve nuclei: see (Adams, Shipp, et al., 2013) and Fig. 2. Mechanistically, descending proprioceptive predictions provide a target or *set point* for peripheral reflex arcs – that respond by minimising (proprioceptive) prediction errors.

2.4. Perception and learning

The preceding aspects of active inference are concerned with expectations about hidden states and causes of sensations and how these expectations minimise prediction error. However, exactly the same arguments apply to the parameters of generative models. Model parameters are quantities that do

not change over time and encode causal regularities and associations. Formally, these are the parameters of the functions in Fig. 2. Neurobiologically, the parameters of a generative model are thought to be encoded by synaptic connection strengths and their Bayesian updates look very much like (experience-dependent) associative plasticity that mediates short and long-term changes in synaptic connectivity: see the Appendix and (Friston, 2008) for details. The distinction between states and parameters of generative models induces the distinction between perceptual inference and learning that proceed over different timescales. Because parameters do not change with time, they accumulate prediction errors over time and are therefore updated at a slower timescale. Equipped with a formal description of perceptual inference and learning, we can now examine the nature of communication using simulations for any given generative model. In the next section, we will apply the computational scheme above to a model of auditory exchange between two systems that are actively trying to infer each other.

3. Birdsong and attractors

This section introduces the simulations of birdsong that we will use to illustrate active inference and communication in subsequent sections. The basic idea here is that the environment unfolds as an ordered sequence of states, whose equations of motion induce attractor manifolds that contain sensory trajectories. If we consider the brain has a generative model of these trajectories, then we would expect to see attractors in neuronal dynamics that are trying to predict sensory input. This form of generative model has a number of plausible characteristics:

Models based upon attractors can generate and therefore encode structured sequences of events, as states flow over different parts of the attractor manifold (a subset of states to which the flow is attracted). These sequences can be simple, such as the quasi-periodic attractors of central pattern generators or can exhibit complicated sequences of the sort associated with itinerant dynamics (Breakspear & Stam, 2005; Rabinovich, Huerta, & Laurent, 2008). Furthermore, hierarchically deployed attractors enable the brain to predict or represent sequences of sequences. This is because any low-level attractor embodies a family of trajectories. A natural example here would be language (Jackendoff, 2002). This means it is possible to generate and represent sequences of sequences and, by induction sequences of sequences of sequences etc. This rests upon the states of neuronal attractors at any cortical level providing control parameters for attractor dynamics at the level below (Kiebel, von Kriegstein, Daunizeau, & Friston, 2009). In the example below, we will show how attractor dynamics furnish generative models of sensory input, which behave much like real brains, when measured electrophysiologically.

We first reproduce the simulations reported in (Friston and Frith, 2015) to illustrate the basic nature of the generalised synchrony induced by predictive coding. We then present new results showing that the acquisition of this synchrony emerges naturally, when two predictive coding schemes try to predict each other. This acquisition resolves the hermeneutic

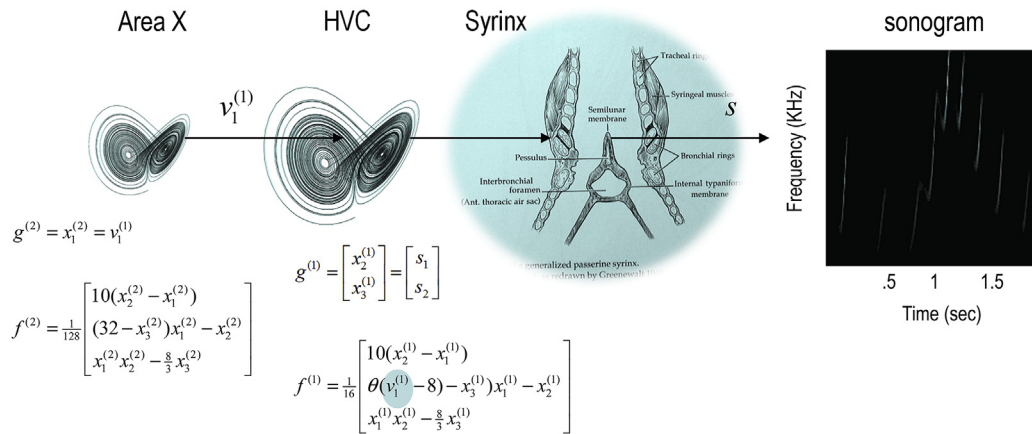


Fig. 3 – Schematic showing the construction of a generative model for birdsongs. The upper panel illustrates the generative model that comprises two Lorenz attractors, where the higher attractor delivers a control parameter (cyan circle) to a lower level attractor, which, in turn, controls a synthetic syrinx to produce amplitude and frequency modulated stimuli. This stimulus is represented as a sonogram in the right panel. The equations represent the hierarchical dynamic model in the form described in the Appendix.

problem by underwriting online inference about the causes of shared sensory consequences. A more detailed description of the birdsong model can be found in (Friston & Kiebel, 2009).

3.1. A synthetic songbird

The example used here deals with the generation and recognition of birdsongs. We imagine that birdsongs are produced by two time-varying states that control the frequency and amplitude of vibrations of the syrinx of a songbird (Fig. 3). There is an extensive modelling effort using attractor models to understand the generation of birdsong at the biomechanical level (Mindlin & Laje, 2005). Here, we use attractors to provide time-varying control over the resulting sonograms. We drive the syrinx with two states of a Lorenz attractor, one controlling the frequency (between two to five KHz) and the other controlling the amplitude or volume. The parameters of the Lorenz attractor were chosen to generate a short sequence of chirps every few hundred milliseconds or so. The Lorenz form for these dynamics is a somewhat arbitrary choice but provides a ubiquitous model of chaotic dynamics in the physical (Poland, 1993) and biological (de Boer & Perelson, 1991) sciences.

To give the generative model a hierarchical structure, we placed a second Lorenz attractor, whose dynamics were an order of magnitude slower, over the first. The first state of the slow (extrasensory) attractor provided a control parameter for the fast (sensory) attractor generating the sonogram. In fluid dynamics, this control parameter is known as a *Rayleigh number* and reflects the degree of convective or turbulent flow, which we will associate with dynamical prosody. In other words, the state of the slower attractor changes the manifold of the fast attractor. This manifold could range from a fixed-point attractor, where the states collapse to zero; through to quasi-periodic and chaotic behaviour associated with a high Rayleigh number. Because higher states

evolve more slowly, they modulate the chaotic behaviour of the sensory attractor, generating songs, where each song comprises a series of distinct chirps. As shown in Fig. 3, the Rayleigh number linking hierarchical levels depends on a model parameter θ that controls the influence of the higher attractor over the lower attractor. High values of this parameter increase the dynamical prosody of the song (inducing successive bifurcations: see Fig. 4). We will use this parameter later to demonstrate perceptual learning and closure of the hermeneutic circle.

3.2. Omission and violation of predictions

To illustrate the predictive nature of predictive coding, perceptual inference was simulated by integrating the above scheme. Formal details of the generative model and integration scheme are provided in the equations in Fig. 3 and the Appendix respectively.¹ A more detailed description of this simulation can be found in a companion paper (Friston and Frith, 2015). In brief, a sonogram was produced using the above composition of Lorenz attractors (with $\theta = 1$) and played to a synthetic bird – who tried to infer the underlying hidden states of the sensory and extrasensory attractors (associated with the HVC and area X respectively). Crucially, we presented two songs to the bird, with and without the final chirps. The corresponding sonograms and percepts (predictions) are shown with their prediction errors in Fig. 5.

The left panels show the stimulus and percept, while the right panels show the stimulus and responses to omission of

¹ For people interested in the technical details – and exploring the effects of changing parameters – the routines generating the figures in this paper are available as part of the academic SPM freeware (<http://www.fil.ion.ucl.ac.uk/spm/>). Annotated Matlab routines can be run and edited by invoking a graphical user interface (by typing DEM) and selecting *Birdsong duet* (DEM_demo_duet.m).

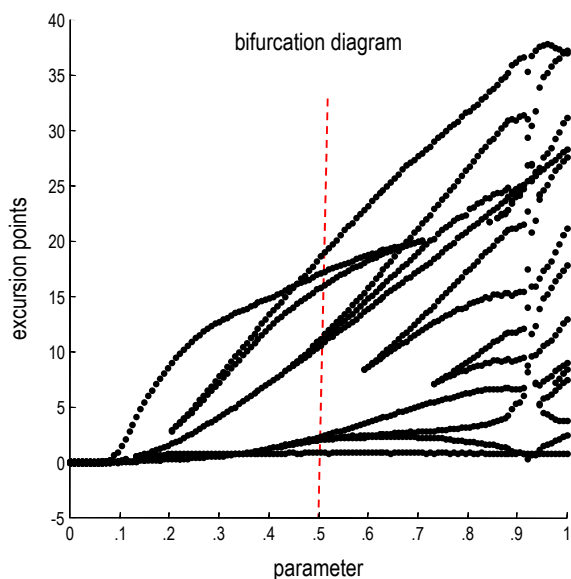


Fig. 4 – Bifurcation diagram showing the dependency of the dynamics on the model parameter. This diagram plots the maxim and minima of a bird's auditory expectations (of hidden sensory states) over 1 sec of simulated time (while singing to itself). This was repeated for 128 values of the parameter ranging from 0 to 1. The ensuing lines show a succession of bifurcations, indicating the emergence of new peaks or transients in the song trajectory. As the parameter falls to zero, the first level attractor is effectively disconnected from the higher attractor and ceases to show chaotic behaviour. The vertical red line shows the value of the parameter we will use later to illustrate learning.

the last chirps. These results illustrate two important phenomena. First, there is a vigorous expression of prediction error after the song terminates prematurely. This reflects the dynamical nature of inference because, at this point, there is no sensory input to predict. In other words, the prediction error is generated entirely by descending predictions. It can be seen that this prediction error (with a percept but no stimulus) is larger than the prediction error associated with the third and fourth chirps that are not perceived (stimulus but no percept). Second, there is a transient percept when the omitted chirp should have occurred. As noted in (Friston and Frith, 2015), this simulation of omission-related responses as measured with ERPs (Bendixen, SanMiguel, & Schröger, 2012) is particularly interesting, given that non-invasive electromagnetic signals arise largely from superficial pyramidal cells. These are the cells thought to encode prediction error (Bastos et al., 2012).

4. A Duet for one

We now turn to the perceptual coupling or communication by simulating two birds that can hear themselves (and each other). Each bird listened for 2 sec, with a low proprioceptive

and a high exteroceptive precision (log-precisions of -8 and 2 respectively) and then sang for 2 sec, with high proprioceptive and attenuated auditory precision (log-precisions of 0 and -2 respectively). The log-precisions at the higher level were fixed at 4 . Crucially, when one bird was singing the other was listening. This turn taking (Wilson & Wilson, 2005) is a natural consequence of active inference because attending to the consequences of action interferes with descending predictions about movement – predictions that are generated before their consequences are evident. There are many examples of the implicit sensory attenuation during self-made acts (Hughes, Desantis, & Waszak, 2013). A physiological illustration can be found in (Agnew, McGettigan, Banks, & Scott, 2013), who show that articulatory movements attenuate auditory responses to speech using fMRI. Furthermore, recent TMS studies suggest that the motor system plays an explicit role in semantic comprehension (Schomers, Kirilina, Weigand, Bajbouj, & Pulvermuller, 2014). We illustrated the importance of sensory attenuation in the context of simulated birdsong in (Friston and Frith, 2015) by showing sensorimotor delays – implicit in articulating a song – destroy the songs dynamical prosody. This means, the birds can only listen or sing but not do both at the same time.

We initialised the simulations with random expectations. This meant that if the birds cannot hear each other, the chaotic dynamics implicit in their generative models causes their expectations to follow divergent trajectories, as shown in Fig. 6. However, if we move the birds within earshot, so that they can hear each other, they synchronise almost immediately. See Fig. 7. This is because the listening bird is immediately entrained by the singing bird to correctly infer the hidden (dynamical) states generating sensations. At the end of the first period of listening, the posterior expectations of both birds approach identical synchrony, which enables the listening bird to take up the song, following on from where the other bird left off. This process has many of the hallmarks of *interactive alignment* in the context of joint action and dialogue (Garrod & Pickering, 2009).

Note that the successive epochs of song are not identical. In other words, the birds are not simply repeating what they have heard – they are pursuing a narrative prescribed by the dynamical attractors (central pattern generators) in their generative models that have been synchronised through sensory exchange. The example in Fig. 7 highlights the fact that the songs articulated by both birds have a rich dynamical vocabulary (including frequency glides, low frequency warbles and amplitude modulated chirps) that is anticipated and reciprocated during the exchange. This means that the birds are singing from the same hymn sheet, preserving sequential and hierarchical structure in their shared narrative. It is this phenomenon – due simply to generalised (in this case identical) synchronisation – that we associate with communication.

Technically speaking, this simulation of communication shows that two identical dynamical systems – that are predicting each other – necessarily show identical synchronisation. In the language of measure-preserving dynamical systems, generalised synchronisation implies the existence of a random dynamical attractor called a *synchronisation manifold*. The synchronisation manifold is just a set of states to which

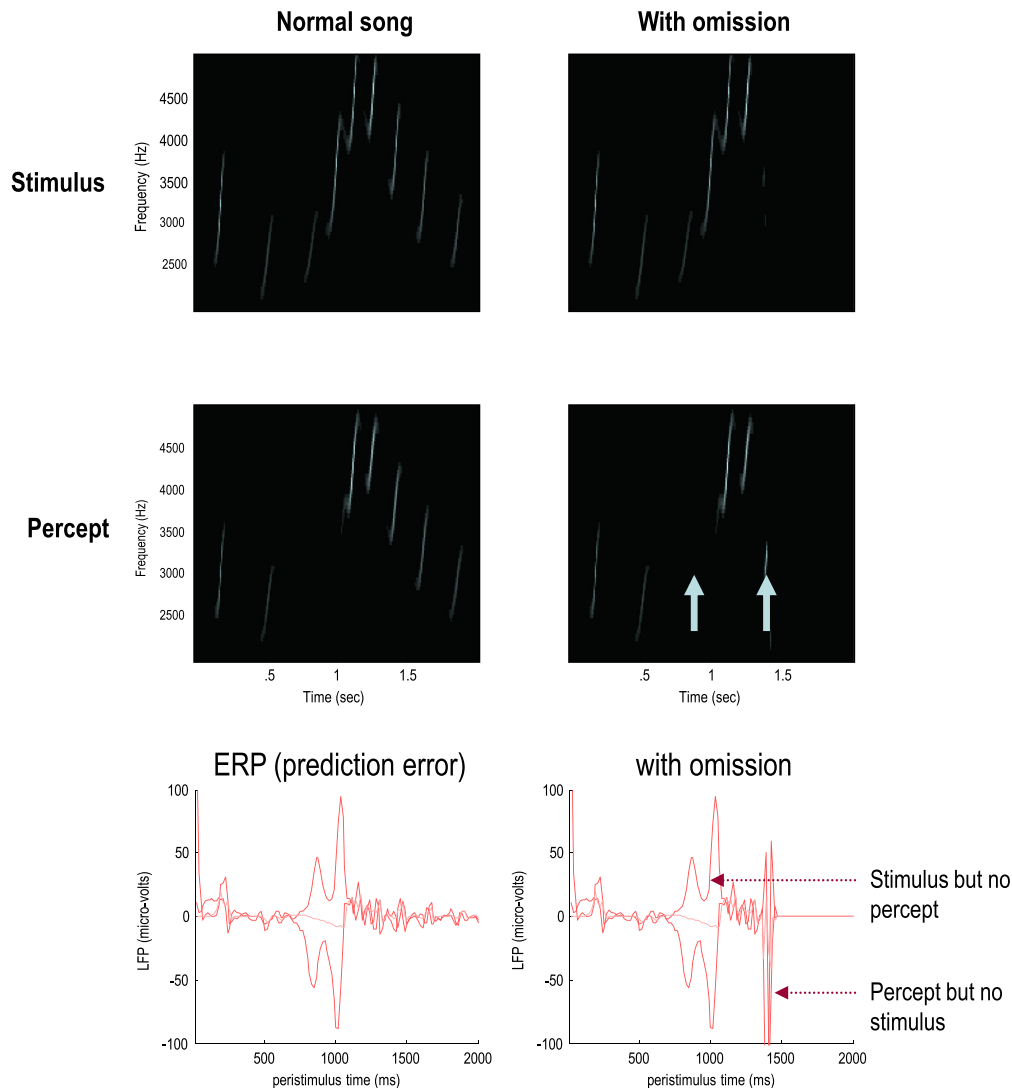


Fig. 5 – Omission-related responses: The left panels show the original song and responses evoked. The right panels show the equivalent responses to omission of the last chirps. The top panels show the stimulus and the middle panels the corresponding percept in sonogram format. The interesting thing to note here is the occurrence of an anomalous percept after termination of the song on the lower right. This corresponds roughly to the chirp that would have been perceived in the absence of omission. The lower panels show the corresponding (precision weighted) prediction error under the two stimuli at both levels. These show a burst of prediction error when a stimulus is missed and at the point that the stimulus is omitted (at times indicated by the arrows on the sonogram). The solid lines correspond to sensory prediction error and the broken lines correspond to extrasensory prediction error at the second level of the generative model.

states are attracted to and thereafter occupy. The simplest example of a synchronisation manifold would be the identity line on a graph plotting homologous states from each system against each other (e.g., the dashed line in Fig. 8). This corresponds to identical synchronisation. Crucially, the attractor (which contains the synchronisation manifold) generally has a low measure or volume. A measure of the attractor's volume is provided by its *measure theoretic entropy* (Sinai, 1959). Although formally distinct from *information theoretic entropy*, both reflect the volume of the attracting set or manifold. This means that minimising prediction errors or free energy (information theoretic entropy) reduces the volume (measure theoretic entropy) of the random dynamical attractor

(synchronisation manifold); thereby inducing generalised synchrony. The synchronisation in the example above is identical because both birds share the same generative model. In the next section, we will illustrate generalised synchronisation using simulations where the birds have different models – and how they learn each other's model to produce identical synchronisation.

4.1. Summary

In summary, these simulations show that generalised synchrony is an emergent property of coupling active inference systems that are trying to predict each other. It is interesting

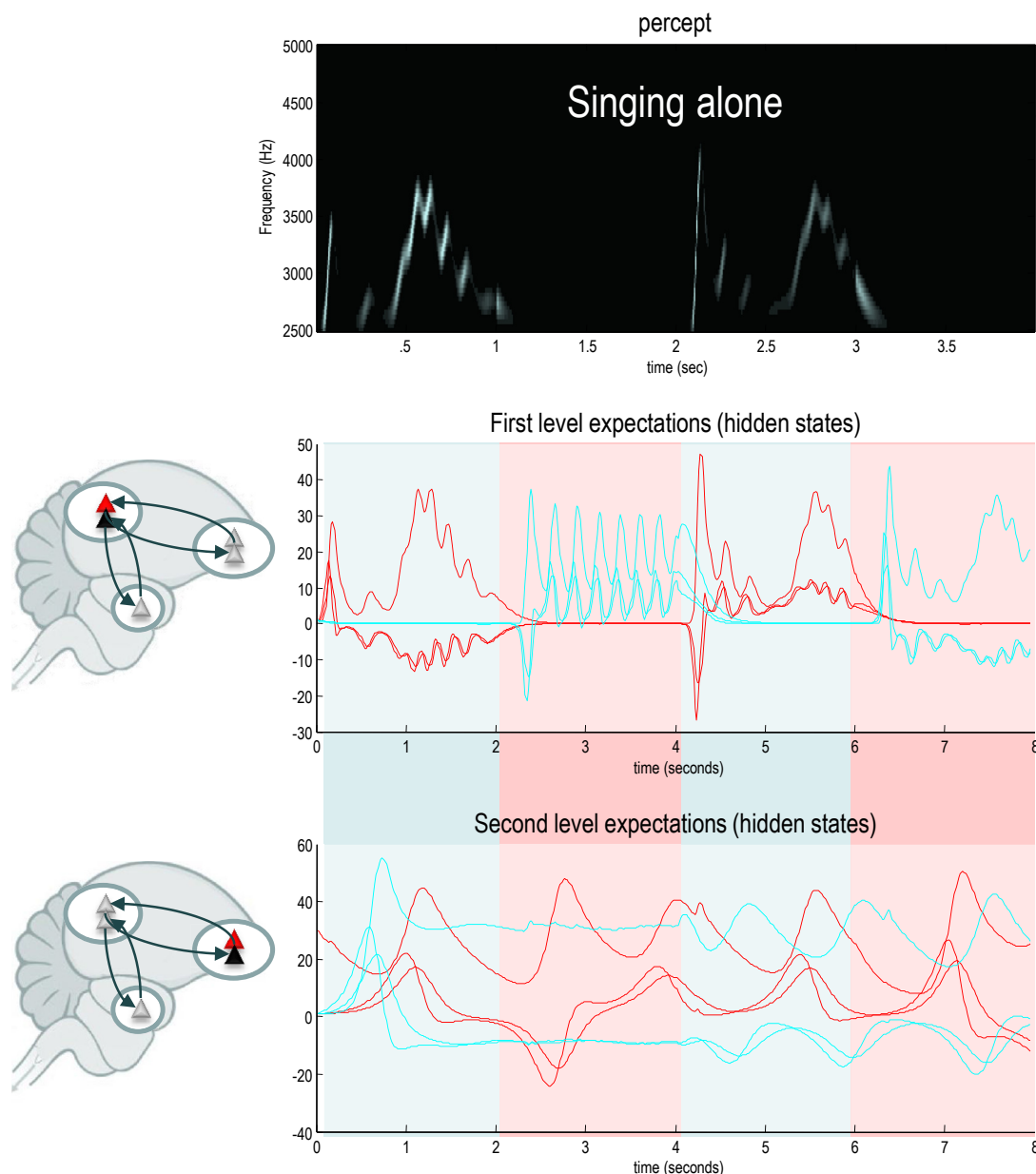


Fig. 6 – A soliloquy for two. In this simulation, two birds with the same generative models – but different initial expectations – sing for 2 sec and then listen for any response. However, the birds cannot hear each other (because they are too far apart) and the successive epochs of songs diverge due to the sensitivity to initial conditions implicit in these (chaotic) generative models. The upper panel shows the sonogram heard by the first (red) bird. Because this bird can only hear itself, the sonogram reflects the proprioceptive predictions based upon posterior expectations in the HVC (middle panel) and area X (lower panel). The posterior expectations for the first bird are shown in red as a function of time – and the equivalent expectations for the second bird are shown in blue. Note that when the birds are listening, their expectations at the first level fall to zero – because they do not hear anything and auditory input is attended (i.e., has a relatively high precision). This does not destroy the slower dynamics in area X, which is able to generate the song again after the end of each listening period. Note also that the second (blue) bird takes a few hundred milliseconds before it starts singing. This is because it takes a little time for the posterior expectations to find the attractor manifold prescribed by the higher level control parameters.

to consider what is being predicted in this context: the sensations of both birds are simply the consequences of some (hierarchically composed and dynamic) hidden states. But what do these states represent? One might argue that they

correspond to some fictive construct that drives the behaviour of one or other bird to produce the sensory consequences that are sampled. But which bird? The sensory consequences are generated, in this setting, by both birds. It therefore seems

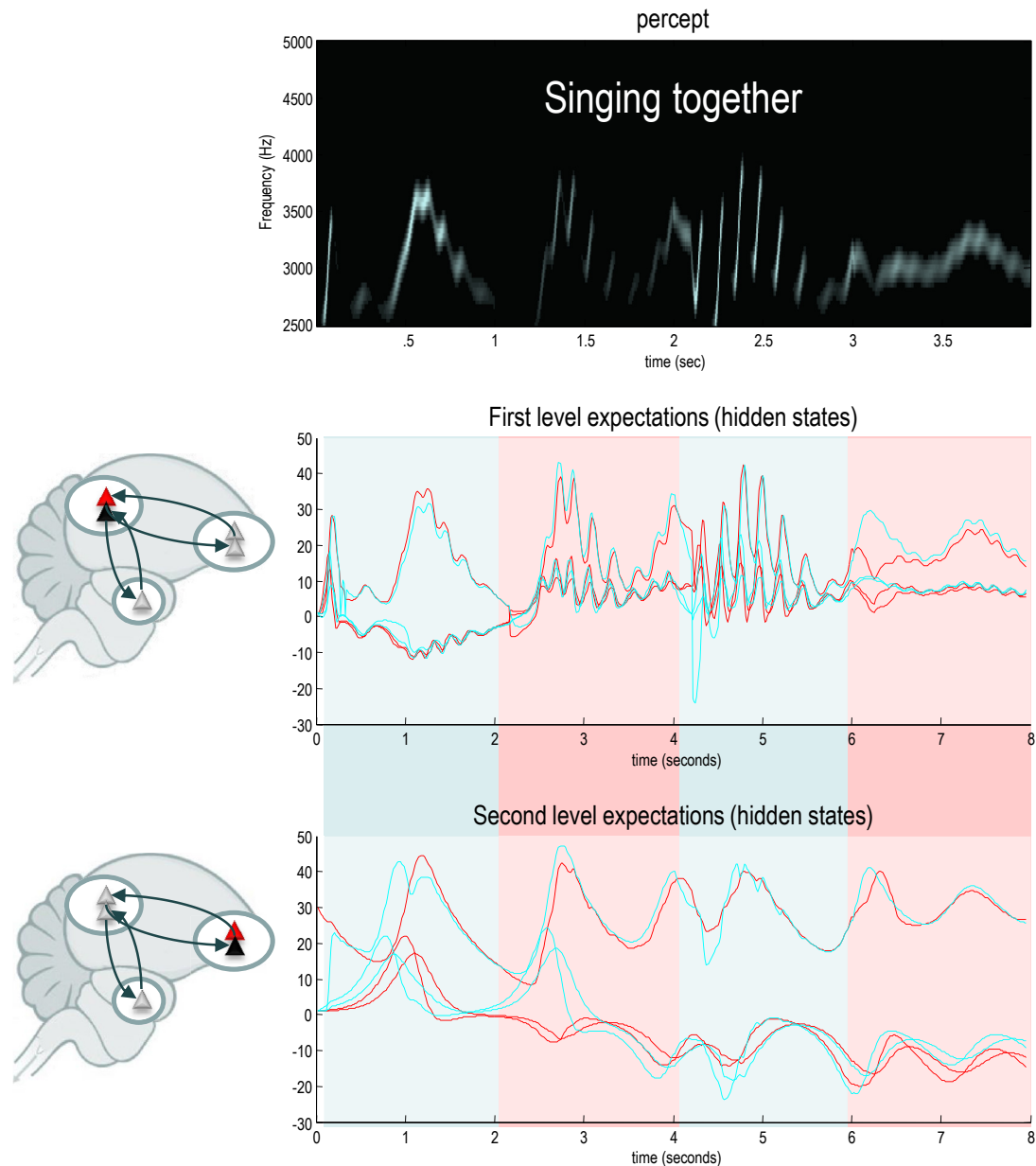


Fig. 7 – Communication and generalised synchrony. This figure uses the same format as Fig. 6; however, here, we have juxtaposed the two birds so that they can hear each other. In this instance, the posterior expectations show identical synchrony at both the sensory and extrasensory hierarchical levels – as shown in the middle and lower panels respectively. Note that the sonogram is continuous over successive 2 sec epochs – being generated alternately by the first and second bird.

plausible to assign these hidden states to both birds and treat the agency as a contextual factor (that depends on sensory attention and attenuation). In other words, from the point of view of one bird, the hidden states are amodal, generating proprioceptive and exteroceptive consequences that are inferred in exactly the same way over time; irrespective of whether sensory consequences are generated by itself or another. The agency or source of sensory consequences is determined not by the hidden states *per se* – but by fluctuations in sensory precision that underlie turn taking. In this sense, the expectations are without agency – they are neither yours nor mine, they are *our* expectations.

5. Closing the hermeneutic cycle

In this final section, we consider how shared narratives emerge as a natural consequence of perceptual learning driven by, and only by, the minimisation of prediction errors or free energy. Using free energy minimisation, [Yildiz, von Kriegstein, and Kiebel \(2013\)](#) have shown that a hierarchy of nonlinear dynamical systems can learn speech samples rapidly and recognize them robustly, even in adverse conditions. Here, we pursue the same theme but when two birds learn from each other; specifically focussing on sensorimotor

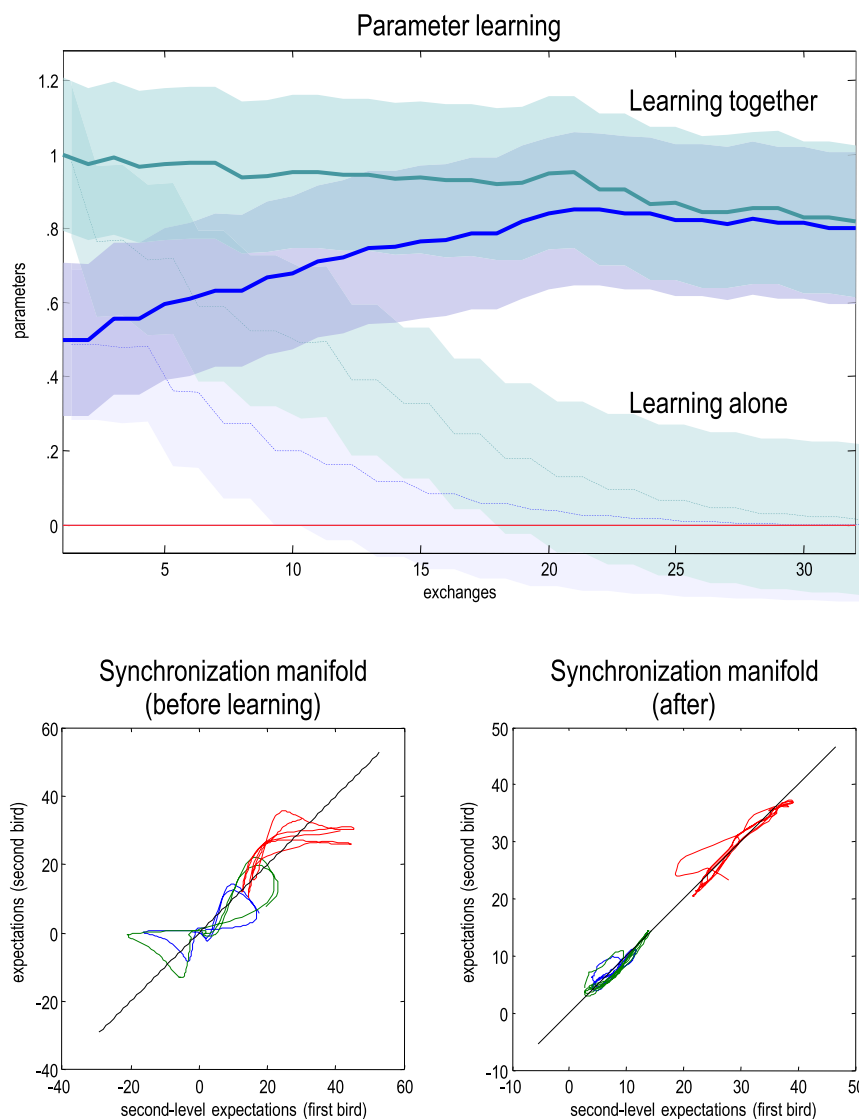


Fig. 8 – Perceptual learning during communication. Upper panel: this shows epoch by epoch changes in the posterior expectations (lines) of the parameter of the first bird (blue) and second bird (green). The shaded areas correspond to 90% (prior) Bayesian confidence intervals. The broken lines (and intervals) report the results of the same simulation but when the birds could not hear each other. The lower panels show the synchronisation of extrasensory (higher) posterior expectations for the first (left panel) and subsequent (right panel) exchanges respectively. This synchronisation is shown by plotting a mixture of expectations (and their temporal derivatives) from the second bird against the equivalent expectations of the first bird, where this mixture is optimised assuming a linear mapping.

learning during which individuals learn to vocalise: see [Bolhuis & Gahr \(2006\)](#) for discussion of birdsong learning. In brief, we repeated the above simulations but reduced the first bird's parameter. This reduces the sensitivity to top-down modulation and suppresses the prosody and richness of the lower attractor. This introduces an asymmetry that precludes identical synchronisation and induces prediction errors that drive learning. This learning corresponds to (epoch by epoch) changes in the posterior expectations of the parameter. In principle, when and only when the parameters are the same is prediction error (free energy) minimised; at which point identical synchronisation should emerge. This effectively

closes the hermeneutic cycle to enable precise communication.

Technically, this sort of perceptual learning is a difficult problem and it took us several tries to find a parameterisation that could be learned efficiently. This is because the inference scheme has to estimate unknown parameters in the context of hidden states that also have to be inferred. Furthermore, the generative model is not only dynamical and nonlinear but chaotic. The example we chose is fairly arbitrary but sufficient to show that a biologically plausible inference scheme can solve this class of learning problem. Our particular parameterisation focuses on the link between the sensory and

extrasensory attractors and can therefore be regarded as a translation of a narrative into a dynamical vocabulary. This means that our two birds start off with a different mapping between (the same) high-level narrative and the way it is articulated. To understand each other, they have to learn each other's mapping (i.e., vocabulary).

In detail, the first bird's parameter was reduced to $\theta = 0.5$, while leaving the second bird's parameter at $\theta = 1$. This means the second bird retained a rich prosody, in relation to the first. After each bird sang to the other, their parameters were updated using a simple form of Bayesian belief updating: the prior expectation was replaced by the posterior expectation (retaining a prior precision of 64). The results of this learning are shown in Fig. 8. The upper panel shows the trajectory of the parameter for both birds over 32 (1 sec) exchanges. It can be seen that the parameters of both birds converge towards each other, so that they both come to articulate a relatively rich sequence of songs. With successive exchanges, the parameter of the first bird (blue line) approaches that of the second bird (green line) and synchronisation emerges (lower right panel). The excursions from the synchronisation manifold during the first exchange highlight the initial difficulties the birds have in predicting each other fluently.

This acquisition and maintenance of this rich discourse can be contrasted with the equivalent learning when the birds cannot hear each other. The resulting trajectories are shown as dashed lines and demonstrate that both birds lose the capacity to maintain their Rayleigh number or dynamical prosody. This is because learning occurs predominantly during listening, when the precision of auditory prediction errors (that drive learning) is not attenuated: note the step-like changes in the parameters. However, when they are listening, each bird only hears silence, which is best explained by an attractor with a Rayleigh number of zero (see Fig. 4). In this situation, both birds emit progressively simpler, low-frequency warbles that disappear by about the 16th exchange (results not shown).

Taken together, these simulations illustrate the circular causality inherent in communicative inference and learning. In other words, one needs to infer or predict a structured sensory exchange before learning can occur; while learning is necessary to render sensory exchanges predictable. Furthermore, although the second bird appears to teach the first bird when they hear each other, both are teaching each other – as evidenced by the fact that the second bird ‘forgot’ how to sing when it could not hear the first. The simulations should not be taken as proof of concept of these points; they simply illustrate the sorts of phenomena that can emerge when dynamical systems are trying to predict and learn from each other.

5.1. Summary

This section has illustrated perceptual learning as a key process in facilitating communication and generalised synchrony in coupled active inference schemes. In this context, prediction errors implicit in the (neurobiologically plausible) predictive coding implementation of active inference are driving

both expectations about hidden states and the learning of model parameters. This means there are two hermeneutic timescales: the first subtending synchronisation and perceptual inference about hidden states of the world and the second slower perceptual learning of the parameters governing fluctuations in the hidden states. The endpoint is a mutual predictability that is underwritten by a convergence of the (parameters of) generative models subtending predictions. Although a simple example, these simulations provide proof of principle that perceptual learning is a sufficient explanation for the emergence of generalised synchrony.

6. Conclusion

The treatment above builds on the ideas introduced by (Friston and Frith, 2015); namely, that generalised synchrony – or synchronisation of chaos – provides a formal metaphor for communication and is a natural consequence of active inference. The contribution of the current paper is to show that the same principle (minimisation of free energy or prediction error), also explains the convergent evolution of hierarchical models that generate mutually sympathetic predictions. We have considered this convergence in terms of perceptual learning that closes the hermeneutic cycle.

There are many issues that we have not considered in the domain of neural hermeneutics and communication. Among these is the possibility that we are equipped with multiple generative models that can be deployed depending upon the situation in which we find ourselves, or the person that we are communicating with. In this instance, the selection of an appropriate generative model – that approximates the model selected by you – would be better understood in terms of Bayesian model selection or averaging (FitzGerald, Dolan, & Friston, 2014). This is an interesting possibility that speaks to inferring the context in which we are communicating and who we are communicating with. This represents another (hierarchical) inference problem. The model selection perspective is interesting because, if correct, it implies a multilateral internal model of how we behave that is entirely dependent upon who we are communicating with. In turn, this begs the question of agency and how it is represented in generative models. For example, if our conceptual narratives are context-sensitive and truly shared, am I the same person (in my mind) when speaking to you, as opposed to somebody else? Furthermore, if I do not have a narrative (or appropriate mapping to a vocabulary) that corresponds to your narrative (or mapping), will I ever be able to communicate with you? This is where perceptual learning comes into its own; enabling the acquisition of new narratives (and vocabularies) or repurposing of existing models; e.g., (Adank, Hagoort, & Bekkering, 2010; Skoe, Krizman, Spitzer, & Kraus, 2014). However – as anyone who has tried to learn a foreign language can testify – learning can be a difficult and slow process, especially for Englishmen like us.

Conflict of interest statement

The authors declare no conflicts of interest.

Acknowledgements

KJF is funded by the Wellcome Trust (088139/Z/09/Z).

Appendix

This brief description of generalized predictive coding is based on (Feldman & Friston, 2010). A more technical description can be found in (Friston, Stephan, Li, & Daunizeau, 2010). This scheme is based on three assumptions:

- The brain minimizes a free energy of sensory inputs defined by a generative model.
- The generative model used by the brain is hierarchical, nonlinear and dynamic.
- Neuronal firing rates encode the expected state of the world, under this model.

Free energy is a quantity from statistics that measures the quality of a model in terms of the probability that it could have generated observed outcomes. This means that minimizing free energy maximizes the Bayesian evidence for the generative model. The second assumption is motivated by noting that the world is both dynamic and nonlinear and that hierarchical causal structure emerges inevitably from a separation of spatial and temporal scales. The final assumption is the Laplace assumption that leads to a simple and flexible neuronal code.

Given these assumptions, one can simulate a whole variety of neuronal processes by specifying the particular equations that constitute the brain's generative model. In brief, these simulations use differential equations that minimize the free energy of sensory input using a generalized gradient descent.

$$\dot{\tilde{\mu}}(t) = D\tilde{\mu}(t) - \partial_{\tilde{\mu}}F(\tilde{s}(t), \tilde{\mu}(t)) \quad \text{A.1}$$

These differential equations say that neuronal activity encoding posterior expectations about (generalized) hidden states of the world $\tilde{\mu} = (\mu, \mu', \mu'', \dots)$ reduce free energy – where free energy $F(\tilde{s}, \tilde{\mu})$ is a function of sensory inputs $\tilde{s} = (s, s', s'', \dots)$ and neuronal activity. This is known as generalized predictive coding or Bayesian filtering. The first term is a prediction based upon a differential matrix operator D that returns the generalized motion of expected hidden states $D\tilde{\mu} = (\mu', \mu'', \mu''', \dots)$. The second (correction) term is usually expressed as a mixture of prediction errors that ensures the changes in posterior expectations are Bayes-optimal predictions about hidden states of the world. To perform neuronal simulations under this scheme, it is only necessary to integrate or solve Equation A.1 to simulate the neuronal dynamics that encode posterior expectations. Posterior expectations depend upon the brain's generative model of the world, which we assume has the following hierarchical form:

$$\begin{aligned} s &= g^{(1)}(x^{(1)}, v^{(1)}) + \exp\left(-\frac{1}{2}\pi_v^{(1)}\right) \cdot \omega_v^{(1)} \\ \dot{x}^{(1)} &= f^{(1)}(x^{(1)}, v^{(1)}) + \exp\left(-\frac{1}{2}\pi_x^{(1)}\right) \cdot \omega_x^{(1)} \\ &\vdots \\ v^{(i-1)} &= g^{(i)}(x^{(i)}, v^{(i)}) + \exp\left(-\frac{1}{2}\pi_v^{(i)}\right) \cdot \omega_v^{(i)} \\ \dot{x}^{(i)} &= f^{(i)}(x^{(i)}, v^{(i)}) + \exp\left(-\frac{1}{2}\pi_x^{(i)}\right) \cdot \omega_x^{(i)} \\ &\vdots \end{aligned} \quad \text{A.2}$$

This equation describes a probability density over the sensory and hidden states that generate sensory input. Here, the hidden states have been divided into hidden states and causes ($x^{(i)}, v^{(i)}$) at the i -th level within the hierarchical model. Hidden states and causes are abstract variables that the brain uses to explain or predict sensations – like the motion of an object in the field of view.

In these models, hidden causes link hierarchical levels, whereas hidden states link dynamics over time. Here, ($f^{(i)}, g^{(i)}$) are nonlinear functions of hidden states and causes that generate hidden causes for the level below and – at the lowest level – sensory inputs. Random fluctuations in the motion of hidden states and causes ($\omega_x^{(i)}, \omega_v^{(i)}$) enter each level of the hierarchy. Gaussian assumptions about these random fluctuations make the model probabilistic. They play the role of sensory noise at the first level and induce uncertainty at higher levels. The amplitudes of these random fluctuations are quantified by their log-precisions ($\pi_x^{(i)}, \pi_v^{(i)}$).

Given the form of the generative model (A.2) we can now write down the differential equations (A.1) describing neuronal dynamics in terms of (precision-weighted) prediction errors. These errors represent the difference between posterior expectations and predicted values, under the generative model (using $A \cdot B \triangleq A^T B$ and omitting higher-order terms):

$$\begin{aligned} \dot{\tilde{\mu}}_x^{(i)} &= D\tilde{\mu}_x^{(i)} + \frac{\partial \tilde{g}^{(i)}}{\partial \tilde{\mu}_x^{(i)}} \cdot \xi_v^{(i)} + \frac{\partial \tilde{f}^{(i)}}{\partial \tilde{\mu}_x^{(i)}} \cdot \xi_x^{(i)} - D^T \xi_x^{(i)} \\ \dot{\tilde{\mu}}_v^{(i)} &= D\tilde{\mu}_v^{(i)} + \frac{\partial \tilde{g}^{(i)}}{\partial \tilde{\mu}_v^{(i)}} \cdot \xi_v^{(i)} + \frac{\partial \tilde{f}^{(i)}}{\partial \tilde{\mu}_v^{(i)}} \cdot \xi_x^{(i)} - \xi_v^{(i+1)} \\ \xi_x^{(i)} &= \exp\left(\tilde{\pi}_x^{(i)}\right) \cdot \tilde{\varepsilon}_x^{(i)} \\ \xi_v^{(i)} &= \exp\left(\tilde{\pi}_v^{(i)}\right) \cdot \tilde{\varepsilon}_v^{(i)} \\ \tilde{\varepsilon}_x^{(i)} &= D\tilde{\mu}_x^{(i)} - \tilde{f}^{(i)}(\tilde{\mu}_x^{(i)}, \tilde{\mu}_v^{(i)}) \\ \tilde{\varepsilon}_v^{(i)} &= \tilde{\mu}_v^{(i-1)} - \tilde{g}^{(i)}(\tilde{\mu}_x^{(i)}, \tilde{\mu}_v^{(i)}) \end{aligned} \quad \text{A.3}$$

This produces a relatively simple update scheme, in which posterior expectations $\tilde{\mu}^{(i)}$ are driven by a mixture of prediction errors $\tilde{\varepsilon}^{(i)}$ that are defined by the equations of the generative model.

In neuronal network terms, Equation A.3 says that error-units compute the difference between expectations at one level and predictions from the level above (where $\xi^{(i)}$ are precision weighted prediction errors at the i -th level of the hierarchy). Conversely, posterior expectations are driven by prediction errors from the same level and the level below. These constitute bottom-up and lateral messages that drive

posterior expectations towards a better prediction to reduce the prediction error in the level below. In neurobiological implementations of this scheme, the sources of bottom-up prediction errors are generally thought to be superficial pyramidal cells, because they send forward (ascending) connections to higher cortical areas. Conversely, predictions are thought to be conveyed from deep pyramidal cells by backward (descending) connections, to target the superficial pyramidal cells encoding prediction error (Bastos et al., 2012; Mumford, 1992).

The corresponding updates for the parameters follow a similar scheme; however, because parameters are time-invariant we have a simpler gradient descent that effectively accumulates free energy gradients over time:

$$\begin{aligned} \dot{\tilde{\mu}}_{\theta}(t) &= -\partial_{\tilde{\mu}} S(\tilde{s}, \tilde{\mu}) \\ \partial_t S &= F(\tilde{s}(t), \tilde{\mu}_x(t), \tilde{\mu}_v(t), \tilde{\mu}_{\theta}) \end{aligned} \quad \text{A.4}$$

Here, $S(\tilde{s}, \tilde{\mu})$ is the anti-derivative of free energy (known as free action). Neurobiologically, when expressed in terms of prediction errors, the solution to Equation A.4 corresponds to associative plasticity. See (Friston, 2008) for details.

REFERENCES

- Adams, R. A., Shipp, S., & Friston, K. J. (2013a). Predictions not commands: active inference in the motor system. *Brain Structure and Function*, 218(3), 611–643.
- Adams, R. A., Stephan, K. E., Brown, H. R., Frith, C. D., & Friston, K. J. (2013b). The computational anatomy of psychosis. *Front Psychiatry*, 4, 47.
- Adank, P., Hagoort, P., & Bekkering, H. (2010). Imitation improves language comprehension. *Psychological Science*, 21(12), 1903–1909.
- Agnew, Z. K., McGettigan, C., Banks, B., & Scott, S. K. (2013). Articulatory movements modulate auditory responses to speech. *NeuroImage*, 73, 191–199.
- Arnal, L. H., & Giraud, A. L. (2012). Cortical oscillations and sensory predictions. *Trends in Cognitive Sciences*, 16(7), 390–398.
- Arnal, L. H., Wyart, V., & Giraud, A. L. (2011). Transitions in neural oscillations reflect prediction errors generated in audiovisual speech. *Nature Neuroscience*, 14(6), 797–801.
- Barreto, E., Josic, K., Morales, C. J., Sander, E., & So, P. (2003). The geometry of chaos synchronization. *Chaos*, 13, 151–164.
- Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., & Friston, K. J. (2012). Canonical microcircuits for predictive coding. *Neuron*, 76(4), 695–711.
- Bendixen, A., SanMiguel, I., & Schröger, E. (2012). Early electrophysiological indicators for predictive processing in audition: a review. *International Journal of Psychophysiology*, 83(2), 120–131.
- de Boer, R. J., & Perelson, A. S. (1991). Size and connectivity as emergent properties of a developing immune network. *Journal of Theoretical Biology*, 149(3), 381–424.
- Bolhuis, J. J., & Gahr, M. (2006). Neural mechanisms of birdsong memory. *Nature Reviews Neuroscience*, 7(5), 347–357.
- Breakspear, M., & Stam, C. J. (2005). Dynamics of a neural system with a multiscale architecture. *Philosophical Transactions of the Royal Society of London B Biological Sciences*, 360(1457), 1051–1074.
- Brown, E., & Brün, M. (2012). The role of prediction in social neuroscience. *Frontiers in Human Neuroscience*, 6, 147.
- Brown, H., Adams, R. A., Parees, I., Edwards, M., & Friston, K. (2013). Active inference, sensory attenuation and illusions. *Cognitive Processing*, 14(4), 411–427.
- Cisek, P. (2007). Cortical mechanisms of action selection: the affordance competition hypothesis. *Philosophical Transactions of the Royal Society of London B Biological Sciences*, 362(1485), 1585–1599.
- Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36(3), 181–204.
- Dayan, P., Hinton, G. E., & Neal, R. (1995). The Helmholtz machine. *Neural Computation*, 7, 889–904.
- Denes, P. B., & Pinson, E. N. (1993). *The speech chain: Physics and biology of spoken language*. Duffield: Worth Publishers.
- Egner, T., & Summerfield, C. (2013). Grounding predictive coding models in empirical neuroscience research. *Behavioral and Brain Sciences*, 36(3), 210–211.
- Feldman, H., & Friston, K. J. (2010). Attention, uncertainty, and free-energy. *Frontiers in Human Neuroscience*, 4, 215.
- FitzGerald, T., Dolan, R., & Friston, K. (2014). Model averaging, optimal inference, and habit formation. *Frontiers in Human Neuroscience*, 8, 457.
- Fletcher, P. C., & Frith, C. D. (2009). Perceiving is believing: a Bayesian approach to explaining the positive symptoms of schizophrenia. *Nature Reviews Neuroscience*, 10(1), 48–58.
- Frank, M. J., Scheres, A., & Sherman, S. J. (2007). Understanding decision-making deficits in neurological conditions: insights from models of natural action selection. *Philosophical Transactions of the Royal Society of London B Biological Sciences*, 362(1485), 1641–1654.
- Friston, K. (2008). Hierarchical models in the brain. *PLoS Computational Biology*, 4(11), e1000211.
- Friston, K., & Frith, C. (2015). A Duet for one. *Consciousness and Cognition*. <http://dx.doi.org/10.1015/j.concog.2015.10.003> (in press).
- Friston, K., Kilner, J., & Harrison, L. (2006). A free energy principle for the brain. *Journal of Physiology Paris*, 100(1–3), 70–87.
- Friston, K., Mattout, J., & Kilner, J. (2011). Action understanding and active inference. *Biological Cybernetics*, 104, 137–160.
- Friston, K., Stephan, K., Li, B., & Daunizeau, J. (2010). Generalised filtering. *Mathematical Problems in Engineering*, 2010, 621670.
- Friston, K. J., & Kiebel, S. J. (2009). Predictive coding under the free-energy principle. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364, 1211–1221.
- Friston, K. J., Shiner, T., Fitzgerald, T., Galea, J. M., Adams, R., Brown, H., et al. (2012). Dopamine, affordance and active inference. *PLoS Computational Biology*, 8(1), e1002327.
- Frith, C., & Wentzer, T. (2013). Neural hermeneutics. In B. Kaldis (Ed.), *Encyclopedia of philosophy and the social sciences*. Thousand Oaks: SAGE publications, Inc.
- Garrod, S., & Pickering, M. (2009). Joint action, interactive alignment, and dialog. *Topics in Cognitive Science*, 1(2), 292–304.
- Gregory, R. L. (1968). Perceptual illusions and brain models. *Proceedings of the Royal Society of London B: Biological Sciences*, 171, 179–196.
- Helmholtz, H. (1866/1962). *Concerning the perceptions in general. Treatise on physiological optics*. New York: Dover. III.
- Hickok, G. (2013). Predictive coding? Yes, but from what source? *Behavioral and Brain Sciences*, 36(4), 358.
- Hohwy, J. (2013). *The predictive mind*. Oxford: Oxford University Press.
- Hughes, G., Desantis, A., & Waszak, F. (2013). Mechanisms of intentional binding and sensory attenuation: the role of temporal prediction, temporal control, identity prediction, and motor prediction. *Psychological Bulletin*, 139(1), 133–151.

- Hunt, B., Ott, E., & Yorke, J. (1997). Differentiable synchronisation of chaos. *Physical Review E*, 55, 4029–4034.
- Huygens, C. (1673). *Horologium oscillatorium*. France: Parisiis.
- Jackendoff, R. (2002). *Foundations of language: Brain, meaning, grammar, evolution*. New York: Oxford Univ. Press.
- Jiang, J., Summerfield, C., & Egner, T. (2013). Attention sharpens the distinction between expected and unexpected percepts in the visual brain. *Journal of Neuroscience*, 33(47), 18438–18447.
- Kiebel, S. J., Daunizeau, J., & Friston, K. (2008). A hierarchy of time-scales and the brain. *PLoS Computational Biology*, 4, e1000209.
- Kiebel, S. J., von Kriegstein, K., Daunizeau, J., & Friston, K. J. (2009). Recognizing sequences of sequences. *PLoS Computational Biology*, 5(8), e1000464.
- Kilner, J. M., Friston, K. J., & Frith, C. D. (2007). Predictive coding: an account of the mirror neuron system. *Cognitive Processing*, 8(3), 159–166.
- Lee, T. S., & Mumford, D. (2003). Hierarchical Bayesian inference in the visual cortex. *Journal of Optical Society of America Optics Image Science and Vision*, 20, 1434–1448.
- Mindlin, G., & Laje, R. (2005). *The physics of birdsong*. Berlin: Springer-Verlag.
- Mumford, D. (1992). On the computational architecture of the neocortex. II. *Biological Cybernetics*, 66, 241–251.
- Nottebohm, F. (2005). The neural basis of birdsong. *PLoS Biology*, 3(5), e164.
- Numminen, J., Salmelin, R., & Hari, R. (1999). Subject's own speech reduces reactivity of the human auditory cortex. *Neuroscience Letters*, 265(2), 119–122.
- Pickering, M. J., & Clark, A. (2014). Getting ahead: forward models and their place in cognitive architecture. *Trends in Cognitive Sciences*, 18(9), 451–456.
- Poland, D. (1993). Cooperative catalysis and chemical chaos: a chemical model for the Lorenz equations. *Physica D*, 65(1), 86–99.
- Rabinovich, M., Huerta, R., & Laurent, G. (2008). Neuroscience. Transient dynamics for neural processing. *Science*, 321(5885), 48–50.
- Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, 2(1), 79–87.
- Schomers, M. R., Kirilina, E., Weigand, A., Bajbouj, M., & Pulvermuller, F. (2014). Causal influence of articulatory motor cortex on comprehending single spoken words: TMS evidence. *Cerebral Cortex*.
- Scott-Phillips, T. C., & Blythe, R. A. (2013). Why is combinatorial communication rare in the natural world, and why is language an exception to this trend? *Journal of the Royal Society Interface*, 10(88), 20130520.
- Sinai, Y. (1959). On the notion of entropy of a dynamical systems. *Doklady of Russian Academy of Science*, 124, 768–771.
- Skoe, E., Krizman, J., Spitzer, E., & Kraus, N. (2014). Prior experience biases subcortical sensitivity to sound patterns. *Journal of Cognitive Neuroscience*, 1–17 [Epub ahead of print].
- Srinivasan, M. V., Laughlin, S. B., & Dubs, A. (1982). Predictive coding: a fresh view of inhibition in the retina. *Proceedings of the Royal Society of London B Biological Sciences*, 216(1205), 427–459.
- Wang, J., Mathalon, D. H., Roach, B. J., Reilly, J., Keedy, S. K., Sweeney, J. A., et al. (2014). Action planning and predictive coding when speaking. *NeuroImage*, 91, 91–98.
- Wilson, M., & Wilson, T. (2005). An oscillator model of the timing of turn-taking. *Psychonomic Bulletin & Review*, 12(6), 957–968.
- Yildiz, I. B., von Kriegstein, K., & Kiebel, S. J. (2013). From birdsong to human speech recognition: bayesian inference on a hierarchy of nonlinear dynamical systems. *PLoS Computational Biology*, 9(9), e1003219.
- Zeki, S., & Shipp, S. (1988). The functional logic of cortical connections. *Nature*, 335, 311–317.