



PAPER

Thinking outside the cortex: social motivation in the evolution and development of language

Supriya Syal and Barbara L. Finlay

Department of Psychology, Cornell University, USA

Abstract

Alteration of the organization of social and motivational neuroanatomical circuitry must have been an essential step in the evolution of human language. Development of vocal communication across species, particularly birdsong, and new research on the neural organization and evolution of social and motivational circuitry, together suggest that human language is the result of an obligatory link of a powerful cortico-striatal learning system, and subcortical socio-motivational circuitry.

Introduction

The essential modification in human evolution enabling language has been addressed for decades. In adults, language is demonstrably dependent on the cortex, so the essential evolutionary modification for language has reasonably been sought in the quality or quantity of cortical computations found there. The crystallized product of a lifelong learning process may not be the best source of insight, however, about what initial alterations permitted it. Language plasticity after early damage raises the most perplexing questions about what in the cortex allows language. The right half of the cortex may be removed, or the left half, and language still develops. The frontal cortex may be removed or severely reduced. Uncomfortably for capacity arguments, total cortical volume may be reduced wholesale by prenatal genetic accident, such as Down's syndrome, or by postnatal accident, and yet language will often survive (Bates, Reilly, Wulfeck, Dronkers, Opie, Fenson, Kiz, Jeffries, Miller & Herbst, 2001; Bates, 2004). An individual may be deaf, blind, or blind and deaf together, and with dedicated tutelage will still learn language with what remains. What conceivable sensory specialization or cortical alteration could survive these assaults?

We will remain agnostic on the cognitive and cortical architecture of language. Instead, we will highlight the changes in human sociality and its motivational correlates, and explore motivation's mechanistic links to language. Of course, unique aspects of human social structure have already been explored widely and at multiple levels. A general coupling between growth of the neocortex and increasing social complexity in primates has been observed (Dunbar & Schultz, 2007). Demon-

strations of how human cultures have solved the problem of stabilization of non-kin altruism, that is, how to stabilize social structures and eliminate cheating such that the manifest benefits of extended altruism to social groups can be had, have been explored in number (Wilson, Van Vugt & O'Gorman, 2008; Bowles, 2006; Richerson & Boyd, 2005; Warneken & Tomasello, 2008; Pagel, 2009). In development, human children, contrasted with young chimpanzees, attend to social cues, share information, join games and generally cooperate, serving a form of social learning seen only in humans (Moll & Tomasello, 2007; see also Preston & de Waal, 2001). What is new, and what we are now able to do is place human social behavior more accurately in a comparative and neuroanatomical context, as a transformation of our understanding of the neural circuitry of sociality and its evolutionary variation has occurred (Newman, 1999; Goodson, 2005) (see Figure 1).

Making explicit this paper's goals: why evolution and development of social communication in non-human animals should directly inform language development and evolution in humans

We will explore the recent advances in the understanding of the social behavior network and its evolutionary variations, highlighting some key findings for those interested in language learning and comparative cognition. We will make some specific proposals about what kind of neuroanatomical circuitry might be changed to link communication with the most fundamental pleasure and reward in human infants. Further, we will hypothesize that such a link constitutes a socio-motivational 'gate' necessary and sufficient for the evolution

Address for correspondence: Supriya Syal, Department of Psychology, Cornell University, Ithaca, NY 14853, USA; e-mail: ss572@cornell.edu

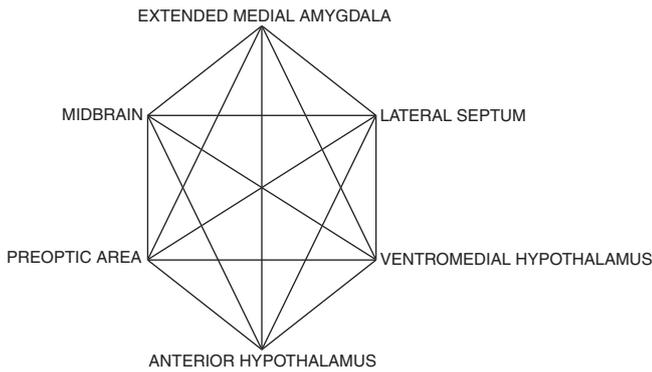


Figure 1 The social behavior network as originally described by Newman, 1999.

of song in a small brain, and perhaps language in a very large one.

The functional neuroanatomy of birdsong will play a central role in this argument. While avian vocal nuclei were first equated with striatal (Edinger, 1988–1903) and then cortical (Jarvis, Gunturkun, Bruce, Csillag, Karten, Kuenzel *et al.*, 2005) mammalian domains, we highlight recent evidence that indicates that much of the avian forebrain is in fact limbic (Medina, 2007; Box 1; Fig-

ure 2) in nature and is thus better equated with mammalian brain regions associated with social motivation. We argue that this better placement of these structures highlights the intrinsic coupling of sight, hearing and vocalization with motivation and underscores the possibility of similar components, largely unexplored, in the neural circuitry of human language learning (Figure 3).

Box 1

A note on developmental neuroanatomy

The vertebrate forebrain is divided into the telencephalon and the diencephalon. The diencephalon contains the regions described by the thalamus and the hypothalamus. The telencephalon is divided into a dorsal pallium and a ventral subpallium. The subpallium develops into the basal ganglia and related structures whereas the pallium forms the cortex and parts of the amygdala. The pallium can be further divided into medial, dorsal, lateral and ventral pallium (Puelles, Kuwana, Puelles, Bulfone, Shimamura, Keleher, Smiga & Rubenstein, 2000). In mammals, the

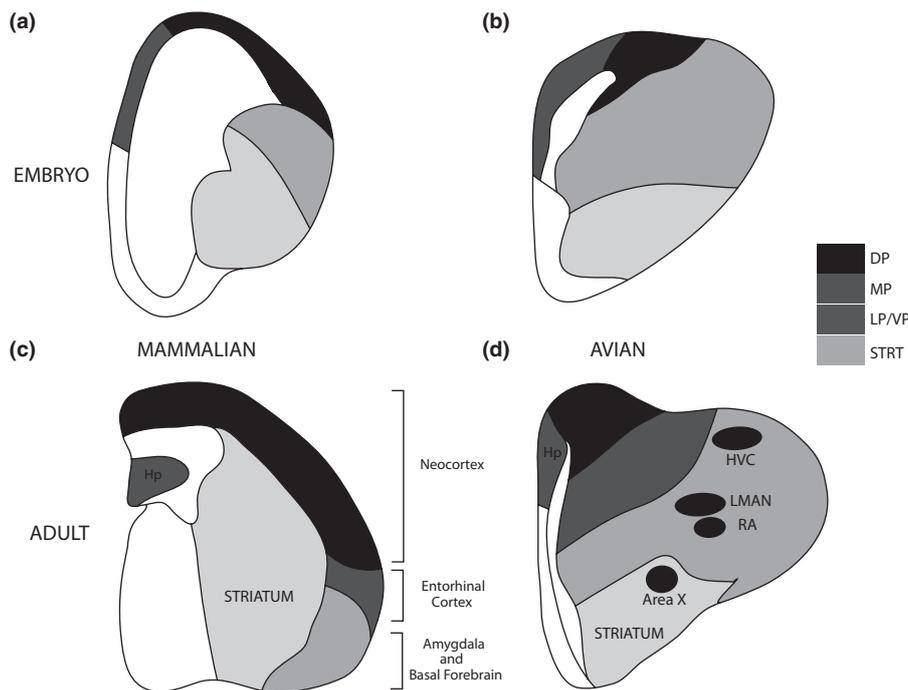


Figure 2 Schematic of transverse sections through mammalian and avian brains at two stages of development. Modified from Figure 2, Medina, 2007. (a) Mammalian embryo, (b) Avian embryo, (c) Mammalian adult, (d) Avian adult with the major vocal nuclei delineated according to embryonic tissue origin. Note that multiple sections along the anterior–posterior axis have been collapsed in order to allow visualization of all vocal nuclei within the same section. Abbreviations: DP – Dorsal Pallium, Hp – Hippocampus, MP – Middle Pallium, LP – Lateral Pallium, VP – Ventral Pallium, STRT – Striatum. Vocal Nuclei: HVC – High Vocal Center, LMAN – Lateral Magnocellular Nucleus of the Nidopallium, RA – Robust Nucleus of the Arcopallium. Only the dorsal pallium gives rise to neo-cortical tissue in mammals (shaded in black). The bird vocal nuclei are not described by this region but instead all lie within tissue derived from lateral and ventral pallia (along with the striatum) which give rise to much of the motivational/social circuitry, i.e. the amygdala and the basal forebrain in mammals. Note that the prevailing view in the field of birdsong still holds that pallial song nuclei are transcribed by mammalian cortical equivalents. We wish to turn attention to the growing body of embryological and topological data that provide emphatic evidence that song nuclei are limbic/motivational in origin (Box 1).

lateral and ventral pallia (originally together referred to as lateral pallium) form the piriform cortex, claustrum and parts of the amygdala. Much of the limbic/motivational circuitry in mammals derives from the lateral and ventral pallia with the medial pallium giving rise to the hippocampal formation. The mammalian isocortex derives from the dorsal pallium. Cross-species comparisons of these subdivisions based on only adult anatomical and functional data have led to the misplaced notion that the latero-ventral pallial regions of the avian brain are homologous to parts of the mammalian isocortex (Reiner, 1993; Karten, 1997), when the only region of the adult bird's brain that the dorsal pallium gives rise to seems to be the 'Wulst' (Medina, 2007). Data derived from embryological origin suggest that most of the avian forebrain becomes the latero-ventral pallium (also called the DVR or the dorsal ventricular ridge) which is homologous to the mammalian claustrum and pallial amygdala (Striedter, 1997; Puelles *et al.*, 2000; Puelles, 2001; Martinez-Garcia, Martinez-Marcos & Lanuza, 2002; Martinez-Garcia, Novejarque & Lanuza, 2007). This would indicate that the nidopallial vocal nuclei of avian species are more soundly positioned in regions homologous to the limbic forebrain of mammals – rather than regions akin to the mammalian neocortex.

The reason this is important is that location often determines function via the class of connections a region makes, the neurotransmitters it expresses, etc. 'Mislocalizing' a lateral pallial derivative (a location which can project directly to hypothalamus and fore-brain, 'motivational' circuitry) as a dorsal pallial structure (which does not have this class of connections) could cause us to misconstrue its functional roles.

We will be discussing far more empirical detail in evolutionary biology and comparative neuroanatomy than is typical for a developmental science audience. In order to win the reader's patience for this enterprise we argue that new work in development and evolution, 'evo-devo', has profoundly changed our understanding of the relationship between the behaviors of existing animals (Kirschner & Gerhart, 2005). What evolutionary pre-suppositions do we bring to understanding roughly 'homologous' behaviors in extant animals, in this case, song in birds versus language in humans? Contrast two scenarios of behavioral evolution, which have at their base two different theories of brain structure and development. In the first, which we will call the 'strict adaptation' model, we imagine an evolutionary random walk from the stem, lizard-like animal to a songbird on the one hand, and a speaking human being on the other. In this scenario, as every feature of the animal's structure and function can be modified in the evolutionary process, the best characterization of any nervous system is of a collection of special-purpose devices (Cosmides & Tooby, 1992) progressively adapted and exapted over the

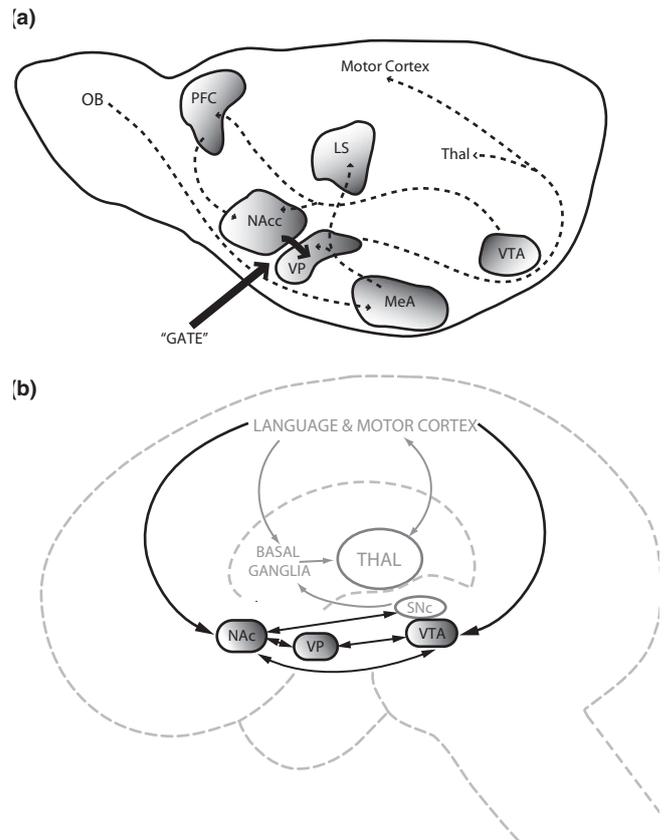


Figure 3 *The social core of language learning: (a) Schematic of a rodent (vole) brain showing the social behavior network. The connection between the Nucleus Accumbens (reward center) and the Ventral Pallidum is highlighted as a potential 'gate' for social motivation whereby the recognition of an individual conspecific becomes potentiated – for example, via the up-regulation of V1aR receptors in monogamous voles (modified from Young & Wang, 2004). (b) Schematic of a human brain showing the basic motor circuitry of language in grey and its interconnectedness with reward centers that might serve similar social gating functions in humans. Abbreviations: LS – Lateral Septum, MeA – Medial Amygdala, NAcc – Nucleus Accumbens, OB – Olfactory Bulb, PFC – Prefrontal Cortex, VP – Ventral Pallidum, VTA – Ventral Tegmental Area, SNc – Substantia Nigra pars caudalis, THAL – Thalamus.*

course of evolutionary time. By the time we compare the nervous systems of the bird and the human, such similarities as we would see would indicate only some distant biases caused by the initially shared nervous system plus the functional similarities that the similar problems of producing and decoding a primarily vocal signal might produce. The functional organization of birdsong might be interesting, but not particularly directly instructive about human language learning.

On the other hand, consider a different scenario which we will call 'developmental conservation', which has arisen from the consistent observation of startling similarities in the basic body and brain plans across vertebrates (Kirschner & Gerhart, 2005). In this scenario, before divergence of birds and mammals, a basic, robust,

and evolvable brain plan had stabilized. Certain kinds of variations, those associated with useful adaptations and niche exploration in the past, were facilitated, and embedded in epigenetic programs that produced stable function in commonly encountered environments. In this scenario, only certain locations in the nervous system can produce vocal behavior, only certain others can link vocal behavior to particular social settings, only certain specific regions can be the site of long-term associative learning. In the 'developmental conservation' model, the functional similarities of song and language are amplified by a conserved basic brain plan and epigenetic operations stabilized over time to facilitate only certain kinds of operations and adaptations. Much debate remains about just how much evolution resembles the first versus the second scenario, and we do not pretend that the issue is settled! Here, however, we are making the strongest possible case for 'developmental conservation'. Not only because we are convinced by the increasing evidence for it, but also to model this relatively new argument for the developmental community who are neither evolutionary biologists nor anatomists.

In the following, we will review evidence for motivational dependencies in the learning of social communication. We will also review briefly the neuroanatomy thought to be important in essential motivation, including what has been discovered recently in evolutionary variation in motivational circuits. We will discuss some mechanisms for recognition and learning from caregivers and argue that small biases in early attention and motivation support the multimodal initial recognition of and eventual complex representation of conspecifics and caregivers. We will conclude with specific suggestions about where to look in the developing brain for the structural and physiological changes which seem likely. In sum, we attempt to return attention to language in its ecological context, acquired in the early development of an obligatorily social, gregarious, and often-altruistic species.

Situated, social, motivated vocal behavior

Going back to the beginning, vertebrates who (independent of learning) vocalize to define territories, or announce their species identity and distance, or readiness to mate, or express pain or alarm, obviously do so in an explicit motivational state. Interestingly, even in organisms that don't need to *learn* their species-typical vocalization (Owren, Dieter, Seyfarth & Cheney, 1993; Seyfarth & Cheney, 1986; Winter, Handley, Ploog & Schott, 1973), social context typically modifies and modulates the nature and the content of such vocalizations (Roush & Snowdon, 1994, 1999; Seyfarth & Cheney, 1997). For instance, changes in social structure will cause adult birds to modify their non-song vocalizations in a phenomenon known as *vocal sharing* whereby individuals change their own vocalizations to resemble those of their conspecifics. Such is the case with black-capped

chickadees (*Poecile atricapillus*) that modify the structure of their call at the time of formation of the winter flock, with all birds converging towards a common structure (Nowicki, 1989). Similarly, when the social groups of starlings (*Sturnus vulgaris*) are changed, they change their song type to match that of their companions (Hausberger, Richard-Yris, Henry, Lepage & Schmidt, 1995). Vocal sharing is used not only to initiate and maintain social bonds within units, but also to signal alliance to compete against other units, as in bottlenose dolphins (*Tursiops truncatus*) (Smolker & Pepper, 1999).

Vocal sharing in a changed social context has been seen in various non-human primates as well. Snowdon and Elowson (1994) found that within 10 weeks of being housed together, all the animals from two different groups of pygmy marmosets changed their call structure in the same direction, regardless of reproductive status. Snowdon, Elowson and Roush (1997) also found that mate pairs of pygmy marmosets (*Cebuella pygmaea*), who had stable trill-structure for months before pairing, significantly changed their trill structure in the 6 weeks immediately after pairing and then this changed structure was stable for at least 3 years after pairing. Female squirrel monkeys produce vocalizations known as 'chucks' with individualistic acoustic properties that constitute a 'vocal signature', and are used to preferentially respond to familiar group members, but not to acoustically similar chucks of unfamiliar conspecifics (Smith, Newman, Hoffman & Fetterley, 1982; Biben & Symmes, 1991). These examples illustrate the intrinsic role that the nature of the social context plays in animal vocalizations.

More elaborate vocal learning, defined as 'the ability to modify acoustic and/or syntactic structure of sounds produced, including imitation and improvisation' (Jarvis, 2007), is a phylogenetically more limited phenomenon. Vocal learning has been unequivocally attributed to seven taxa in all, of which three are birds – songbirds, hummingbirds and parrots; and four are mammals – cetaceans, elephants, bats and humans (Jarvis, 2007). In organisms that learn a complex vocal communication system, social context is typically necessary for this process. Deprivation studies were the starting point for this research. Studies of human children (Fromkin, Krashen, Curtis, Rigler & Rigler, 1974; Lane, 1976) or songbirds (Marler, 1970; Thorpe, 1958) that were raised in isolation showed that vocal learning from conspecifics is necessary for the development of normal speech. It is noteworthy that both the rudimentary speech of children and the isolate song of birds reared in a non-social setting, while simplified, do show some species-specific features (Marler & Sherman, 1985). While this is most often cited as evidence for either an innate model for language (Chomsky, 1981; Fodor, 1983) or an innate human capacity for learning to segment and group sensory stimuli (Elman, Bates & Johnson, 1996), there is an interesting way to recast this finding. Perhaps the underdeveloped vocal repertoires that characterize these infants and nestlings show what is obtained when motivation is subtracted from the vocal learning

equation, and thus seriously misrepresent the structure and preferences which both bring to vocal communication.

The social circuit

Until the last ten years, the circuitry of the brain central to social behavior was described as a collection of sub-cortical centers and areas with specific labels, i.e. male aggression, female parental care and so forth. Integrating across an enormous number of studies, Sarah Newman (1999) proposed a new formulation of the 'social network' in the mammalian brain comprising six heavily interconnected limbic system areas with interlaced functions (Figure 1).

In this conception, any given social behavior is an emergent property of the temporal and dynamic pattern of activity across this social network in the brain, replacing the notion that each node is a center for some category of motivated behavior. Rather, varying levels of activity across the nodes signify a given behavior. The sensory stimuli important for these behaviors differ with species, sex, maturational state and context, as the motor manifestations vary with each animal's competence. Ecological and social niches, such as gregarious versus territorial or diurnal versus nocturnal systems, vary as well, and the rather astonishing observation is the phylogenetic stability of the location of this system in the face of such sensory, motor and niche variability (Newman, 1999, and references therein). Also stable are the essential functions of the social circuit: mate choice and mating, parental behavior, territoriality and aggression, affiliation and so forth, while the instantiation of each function in a particular body and ecology varies. This polysensory, distributed nature of the social network may offer a way over evolutionary time to modulate sensory preferences and dispositions with relative ease while conserving core functionality.

Using this framework, a discussion can begin of how such a neural system evolves to produce varied vertebrate social behaviors, how vital functions are paired with the sensory systems each animal best employs, its motor abilities and motivations. There is already evidence to indicate that the dynamic activation of this network can 'code' for the (negative or positive) valence associated with social encounters. Goodson, Evans, Lindberg and Allen (2005) studied a set of closely related finches which varied in their preferred social context, from solitary and territorial to flocking, but had similar mating systems. They found species-specific differences in immediate early gene (IEG) expression in the medial extended amygdala following exposure to a same-sex conspecific. Activity in this region was negatively correlated with species-typical group size – greater activation of the extended amygdala was associated with the response of a territorial species to same-sex conspecifics – a negatively valenced social encounter, while there was less activity if the bird welcomed the 'intruder', as in the case of the gregarious species.

Both behavioral differences and brain activity differences can also be related to neuropeptide receptor distributions. While the brain distributions of the actual neuropeptides such as oxytocin and vasopressin are very highly conserved across species (Wang, Zhou, Hulihan & Insel, 1996), the location and density of the *receptors* for these peptides are not conserved (Young, Nilsen, Waymire, MacGregor & Insel, 1999). The evolutionary alteration of receptor densities has been well described in the case of two closely related voles. Some voles are socially monogamous, sharing parental care; in these voles, copulation with a partner fixes an enduring preference for the partner's company, not shared by closely related promiscuous species. In the monogamous species, expression of a receptor for vasopressin in reward areas of the basal forebrain permits a link to be forged between the specific representation of an individual vole signaled by vasopressin activity, and reward and comfort itself. With the presence of the vasopressin receptor, a gate is opened between the sensory representation of the individual mate and basic reward, stronger than the general links that can be made between any random cue and a positive outcome.

Specifically, in the monogamous prairie vole, vasopressin receptor (V1aR) binding (concentration) is very dense in the ventral forebrain, in particular in the ventral pallidum which is a region associated with reward and motivation (Lim, Murphy & Young, 2004). Injection of vasopressin into the brain, paired with presentation of a female vole, increases the tendency to affiliate with that female, while blocking vasopressin activity abolishes established partner preference (Lim & Young, 2002, 2004). On the other hand, non-monogamous montane and meadow voles naturally have very low densities of V1aR receptors in this region (Insel, Wang & Ferris, 1994) but experimentally increasing the number of these receptors can lead to pair bond formation even in these promiscuous species (Lim, Hammock & Young, 2004).

Differences in neuropeptide activity within the social network also underlie other kinds of social preferences in birds. Goodson and Evans (2004) studied the social network in a range of species of birds that differed in their coloniality. Different levels of colocalization of AVT (Arginine VasoTocin – the avian orthologue of mammalian vasopressin) and immediate early gene expression in the medial bed nucleus of the stria terminalis (BSTm) and lateral septum marked the varying evaluation of particular social situations by bird species with different levels of gregariousness. The design of these experiments did not permit overt social behavior, thus highlighting evaluation and motivation rather than action. In essence these data indicate that presenting a bird with stimuli that are positively valenced within its ecological niche causes an increase in uptake of AVT, and immediate early gene activity, whereas the opposite is true for negatively valenced stimuli. So for instance, to a gregarious bird, both partners and same-sex conspecifics present opportunities for rewarding social interaction and a corresponding increase in AVT-IEG colocalization is

seen in these birds upon presentation of either of those classes of stimuli. To a territorial bird, on the other hand, a same-sex conspecific signals a potential antagonistic encounter – and this is associated with a decrease in AVTir-IEG colocalization, whereas the presence of the subject bird's partner – presumably a positive stimulus – causes an increase in colocalization.

Whether the effects of vasopressin occur directly through association with reward pathways in the ventral pallidum or through the neuroendocrine mechanisms associated with stress and social anxiety (Goodson & Evans, 2004), the activity of this neuropeptide seems to provide a robust system under which the evolution of sociality can be probed. In effect, one can think of the receptor distribution and concentration of vasopressin receptors as locks that can migrate and vary in number, with keys always standing ready. In experiments, changing the level of only one receptor appears able to produce a change in social preference enough to alter a species' life history and species-level social structure. Vasopressin receptor levels alone, of course, probably are not the only distinctions between social species and non-social species over an evolutionary time-scale. The relative simplicity of the structural changes necessary to bring about large-scale changes in the ecological niche of a given organism is, however, impressive, and candidate genes have already been described in the case of the voles (Young *et al.*, 1999; Hammock & Young, 2004, 2005).

Obscure objects of desire: representing an individual

Before proceeding, it is worthwhile to stand back and consider some comparative information on what 'specific representation of an individual' may entail. Early findings in ethology characterized interactions between animals as specially evolved stimulus–response linkages, such as herring gulls pecking at the red spot on the feeder's beak (Alcock, 2009). These early examples of stimulus–response singularities do not characterize the representational complexity of most vertebrate social interactions particularly well. The research that arose from these first examples showed that a large range of mammals (and probably most vertebrates) come to be endowed with a far more robust representation of a conspecific than a single sensory emblem. The corpus of work by Eric Keverne and colleagues on maternal imprinting on lambs is instructive (Broad, Curley & Keverne 2006). Flocks of sheep give birth relatively simultaneously in the spring, and the ewe has a specialized, time-limited postnatal learning period, hormone dependent, in which she will learn to recognize her own lamb. Allowing the lamb at first to suckle is dependent on olfaction, but the ewe must solve the greater problem of recognizing her own lamb amongst others within a day or so. While olfaction can be involved in this, it is not essential, and all relevant modalities are brought to bear: the bleat, the face and body appearance. The object of

maternal desire is not a value in a single sensory channel, but an individual lamb. Similarly in rodents, the various separable odorants associated with an individual, encountered in different contexts, serve as tokens of the whole (Johnston, 2005). The social behavior network may be seen as the conserved neural structure that assembles on the one hand relevant sensory dimensions into a representation of an individual (either by innate bias or experience-expectant learning), and on the other attaches that representation to the motivations and actions appropriate to their social roles: parent, child, mate or rival. With situation, and maturation, these attachments also change as required.

Turning back to language evolution, we can use this framework to better understand how, even in closely related species, the motivational significance of a particular conspecific role could differ radically, depending on which gates are opened to reward and comfort, or to anxiety and aversion. These simple but significant changes can immediately alter the entire social setting and preferences of an individual, who they attend to and stay close to. The increased social orientation of human infants to solicit other humans' play and help, compared to chimpanzees, is exactly the type of change for which we are attempting to provide a mechanistic account (Tomasello, 1999). The utility of such changes to motivate language learning seems reasonable, but social circuitry *per se* as we have described it so far carries no explicit way to potentiate particular learned behaviors.

Links between the social circuit and motivation systems

A mountain of evidence exists in support of the claim that limbic structures (and the hypothalamus) code for motivational processes, and the reader is referred to many excellent reviews for further reading (Mogenson, Jones & Yim, 1980; Cardinal, Parkinson, Hall & Everitt, 2002). For the purposes of this essay, we wish to highlight the overlap between much of the social circuit and parts of the motivation/emotion circuits (Figures 2 and 3). In the large expanse of the limbic forebrain, the extended amygdala forms a relay center for information crossover between the cortical and brainstem regions. While it acts in concert with other regions of the brain in encoding emotional processes, the amygdala, via its modulation of both attentional and motivational arousal, is one very salient functional seat of motivation in the mammalian brain. The connections between the amygdala and the ventral tegmental area, via the nucleus accumbens – a structure located in the ventral striatum – form a large part of the mesolimbic dopamine pathway that modulates behavioral responses to stimuli that activate feelings of reward or motivation. According to the Depue and Morrone-Strupinsky (2005) model of affiliative bonding in humans, the convergence of dopaminergic, neuropeptide and opiate systems in the ventral striatum

provides for the human capacity to feel rewarded by affiliative (social) stimuli.

While additional regions of the mammalian brain have been associated with motivation-related processes (hippocampal formation, the cingulate cortex, and limbic midbrain areas; Morgane, Galler & Mokler, 2005), one that bears mention here is the prefrontal cortex (PFC). The prefrontal cortex regulates complex memory, attention and cognitive processes and is also an integral part of the motivation circuit via its extensive bi-directional connectivity with the limbic forebrain (Damasio, Grabowski, Frank, Galaburda & Damasio, 1994; Goldman-Rakic, 1999). In humans, increased gray matter density in the orbitofrontal cortex (PFC) and the ventral striatum (Lebreton, Barnes, Miettunen, Peltonen, Ridler, Veijola *et al.*, 2009) and increased connectivity between the frontal cortex (including the orbitofrontal cortex) and striatum (Cohen, Schoene-Bake, Elger & Weber, 2009) are associated with increased responsiveness to socially defined reward and increased disposition to social relationships and attachments.

So how do these structures tie into vocal learning and/or production? It is instructive to think of vocal learning as a specialized learning behavior that, like most (if not all) behavioral systems that incorporate any kind of learning, exists in a motivational context. The intimate linkages between the motivation and social circuits provide us with a system that can represent and attach motivational value to social conspecifics. All that is required, over evolutionary time, is the linkage of vocal centers to that social-motivation system – and together these will comprise the vocal learning system (Figure 3a and 3b). And these links are found at both the anatomical level and the behavioral level. For instance, projections of the ventral tegmental area to the avian vocal centers modulate early gene activity related to directed or undirected singing in zebra finches (Hara, Kubikova, Hessler & Jarvis, 2007). The extensive work on birdsong, and the fairly recent re-understanding of the homologies between the neuroanatomical regions involved in vocal communication in birds and in humans, will be our vehicle to a mechanistic understanding of where brain changes may have occurred.

Birdsong and human language

Many parallels between birdsong and human speech have been explored; perhaps the most striking is the interplay between production and perception in the development and maintenance of vocal communication (Doupe & Kuhl, 1999). The parallels in comparative neuroanatomy, critical periods, mechanisms of learning, lateralization and so on are also striking. From this voluminous literature, we will concentrate on those studies that emphasize the natural ecology of learning, its neuroanatomical motivational substrates and their interaction.

Although a full discussion is beyond the scope of this review, the renaming of the structures of the bird telen-

cephalon according to their proper site of origin, neuropharmacology and connectivity to allow proper homologies to other vertebrate telencephalon (including our own) has brought many 'birdsong nuclei' into the basal forebrain motivational circuitry where formerly they had been given quasi-cortical status. Under this scheme (Medina, 2007; Box 1; Figure 2), a number of key nuclei in the song control system now fall in the region of brain homologous to basal forebrain and amygdala. In Figure 2a–d, the divisions of the embryonic telencephalon which give rise to the hippocampus, dorsal pallium and isocortex; nidopallium, entorhinal cortex and amygdala; arcopallium and the striatum; as well as the thalamic-hypothalamic core are given. Looking at the location of birdsong nuclei in this light, it is clear that most fall into those regions associated with direct access to motivational and social circuitry, and not the neocortex. According to the argument we made earlier about the essential conservation of fundamental operations in the vertebrate brain, if much of the circuitry of this learned, social, vocal behavior falls in the extended amygdala and associated limbic regions in birds, we should seek its homologue(s) in the other vertebrate vocal learner we commonly study, the human infant.

In 1969, Immelmann conducted what is now held to be a seminal experiment in the progression of the science of birdsong learning. He cross-fostered infant zebra finches in Bengalese finch families and found that the zebra finches learned the song of their heterospecific fathers even when adult conspecific males could be heard nearby (Immelmann, 1969). Since then, zebra finches have been also shown to need social interaction with a real-bird tutor (most often, the father) in order to learn a song from him and typically do not learn well from tapes (Eales, 1989; Mann & Slater, 1995). If, however, the experimental setting is such that a subject bird is required to activate a tape by pressing a key to elicit the zebra finch song model, the song is learnt. Yoked controls who can hear the same playback, however, do not learn the song (Adret, 1993). Similarly, parrots who are trained to produce a variety of words and labels in human language do not learn through watching videotapes or listening to audio tapes or by watching non-interactive humans (Pepperberg, 1999). The commonality between a live tutor and the controllable playback is response contingency. While the results of the Adret, 1993, study are contentious because a subsequent study was not able to replicate these specific effects in zebra finches (Houx & ten Cate, 1999), there is now ample evidence from human studies to indicate that response contingency is crucial to infant learning of advanced vocal forms (Bloom, Russel & Wessenberg, 1987; Goldstein & Schwade, 2008). In addition, response contingency in the form of turn-taking interactions is an early feature of parent–infant face-to-face interactions (Papousek, Papousek & Bornstein, 1985) and human infants respond to social contingency from a very young age and show sensitivity to familiar contingency levels based on the responsiveness of their

caregivers (Bigelow, 1998; Bigelow & Rochat, 2006). It has been suggested that social contingency aids learning by focusing infant attention on relevant features of stimulus information (Kuhl, 2007). It is thus a property of the interaction itself, and not sensory signals or motor output that gates learning, an observation that is in fact fairly pervasive.

In zebra finches, neural activity during singing is strongly modulated by the presence of a listener. Both the magnitude and the variability of activity in the LMAN (lateral magnocellular nucleus of the nidopallium, song nucleus) and Area X (basal ganglia, song nucleus) is dependent on the presence of a conspecific – the activity is lower and more uniform during directed singing (to a female) and higher and more variable during undirected singing (alone) (Hessler & Doupe, 1999). Dopamine levels in Area X are significantly higher during directed singing as compared to undirected singing – possibly because of higher dopamine release from the VTA-SNc (Ventral tegmental area-Substantia nigra pars caudalis, reward region) terminal that project to Area X (Sasaki, Sotnikova, Gainetdinov & Jarvis, 2006). Area X is necessary for song learning and is part of the AFP (anterior forebrain pathway: cortex → basal ganglia → thalamus → cortex), which is involved in auditory feedback-based evaluation and adaptive modification of the bird's own song. High dopamine levels in the striatum of birds during directed song parallel high dopamine levels in the striatum after drug administration in mammals (Sasaki *et al.*, 2006) – pointing to the rewarding nature of the 'feeling of being heard'. Increased glutamate activity in the VTA is also seen after intense social interaction (Huang & Hessler, 2008). The significance of social context in vocal behavior is further highlighted by the observation that new neuronal recruitment to the HVC (high vocal center, song nucleus), Area X and NC (nidopallium caudale, cortex) increases in the brains of adult males who are placed in a complex social context (large aviary setting) as compared to males who are placed in a simple social context (pair cages) (Lipkind, Nottebohm, Rado & Barnea, 2002). This is true even though birds in the simple social environment sing much more than the birds in the complex environment, indicating that the neuronal recruitment is associated with the processing and storage of rich auditory input in a rich social setting rather than mere vocal exertion (Adar, Lotem & Barnea, 2008).

The concept of intrinsic reward in vocalization (Cheng & Durand 2004), with direct physiological outcomes such as ovulation, has been studied extensively in ring doves (*Streptopelia risoria*). In the female ring dove, estrogen acts on a midbrain song nucleus (mICo) to induce a female nest coo in response to a male courtship coo (Cohen & Cheng, 1981). Both auditory and proprioceptive feedback from the female bird's own cooing (and not the male's cooing) initiates an endocrine cascade that ultimately results in egg-laying (Cheng, 2003). In brown-headed cowbirds (*Molothrus ater*), the

female provides feedback to a singing male via a wing-stroke when she likes his song – something that causes him to incorporate more of the 'approved' acoustic forms into his vocal repertoire (West & King, 1988). Such a gesture on the part of the female cowbird could act like the cooing vocalization in ring doves, whereby the wing stroke, acting via the circuit that receives modulatory affective input from the arcopallium/amygdala, could not only be a signal of the willingness of the bird to copulate, but in fact self-stimulate her to do so.

In summary, the extensive and expansive connectivity that exists between motivation systems and the vocal nuclei in avian species, along with behavioral data that implicate affective systems in vocal learning processes, suggests that neural pathways that mediate emotional and motivational states and vocal control pathways form an *integrated* system of communication in birds. Further, these pathways receive constant direction from 'social road-signs' within the birds' ecological niche. In the avian species that display learned vocal behavior, the learning environment is best viewed as a consolidated system that incorporates the entire infant bird, its tutor, the interaction between the two, and the effect of each actor on its own and the other's nervous system.

Social motivation in human language learning

The role of motivation or affect during communicative interactions between human mothers and infants has long been appreciated. Factors such as the prosody of caregivers' early speech to infants, along with facial movements, form non-linguistic affective behaviors that 'attract and sustain the infants' attention to what will become a linguistically important stream of cues' (Locke, 1993). Caregivers use infant-directed speech to facilitate infant arousal and motivation, which in turn facilitates infant learning of language. The prosody of infant-directed (ID) speech is different from adult-directed speech. Caregivers' speech to infants is often characterized by shorter utterances, longer pauses, higher pitch and wider pitch excursions (Fernald & Simon 1984; Fernald, Taeschner, Dun, Papousek, de Boysson-Bardies & Fukui, 1989). Infants can discriminate affective vocal expressions in ID speech in their own language as well as foreign languages, and the intonation of ID speech can elicit emotional responses from infants (Fernald, 1992a, 1992b). The intonation of ID speech has been shown to effectively engage infant interest (Kuhl, Coffey-Corina, Padden & Dawson, 2005). Finally, mothers reliably use falling pitch contours in speech to infants when trying to soothe (Fernald, Kermanschachi & Lees, 1984) and rising pitch contours when trying to engage infant attention (in a social interaction) (Stern, Spieker & MacKain, 1982). These data have led to the idea that maternal prosody is finely tuned to infant attention and arousal level (Fernald, 1985) and can be used to modulate infant motivational states. In further support of the claim that motivation

drives acquisition of communicative faculties, prelinguistic vocal learning in human infants has also been shown to be modulated by social reinforcement. Contingent responsiveness from caregivers, both silent and vocal, is associated with vocal learning in human infants (Goldstein, King & West, 2003; Goldstein & Schwade, 2008). Further infant learning of phonemic contrasts in a foreign language is contingent on their presentation by a socially interactive live tutor (Kuhl, Tsao & Liu, 2003). Based on this and other work, Kuhl (2007) has advanced the 'social gating' hypothesis whereby infant language is gated by the motivating properties (such as attention and arousal) inherent in social interactions.

In some bilingual households, children acquire proficiency in only one language (Wong Fillmore, 1991). It has been argued that for these children, the language that is not learned is the one that is less 'significant' to the child in that it is used less often in settings wherein the child is likely to be motivated, for example among peer or sibling groups.

Aberrant language acquisition

A particularly interesting case of the interplay between motivational and vocal centers in the brain is autism. Whether or not limbic system dysfunction plays a causal role in the pathology of this disorder remains contentious but there does seem to be some support for the idea that the amygdala and hippocampus of autistic individuals develop abnormally (Schumann, Hamstra, Goodlin-Jones, Lotspeich, Kwon, Buonocore, Lammers, Reiss & Amaral, 2004; Nacewicz, Dalton, Johnstone, Long, McAuliff, Oakes, Alexander & Davidson, 2006). Approximately 25% of all children with autism never develop functional language abilities (Klinger, Dawson & Renner, 2002). Unlike typically developing children, autistic children prefer non-speech samples to acoustically matched samples of ID speech and reportedly display no 'interest' in language, which has been correlated with aberrant neural responses to speech. They also have trouble maintaining attention in linguistic tasks (Kuhl *et al.*, 2005) and these attention deficits in children with autism are also evident in their inability to form and maintain joint focus of attention (Osterling & Dawson, 1994; Osterling, Dawson & Munson, 2002). At around 9 months of age, typically developing human infants start to engage in triadic interactions that involve them, their caregivers and a third external entity that might be a toy or another person or anything in the outside environment. These interactions involve joint attention between the infant and the caregiver and it has been argued that the ability to do so is a necessary milestone in the development of communicative faculties (Tomasello, 1999). In addition to this, overall attention and shared visual attention between infant and tutor predict the degree to which infants will learn the phonemes and words of a language (Conboy, Brooks, Meltzoff & Kuhl, 2008). Autistic adults have an exaggerated anxiety response to

eyes, and when looking at faces, attend to the mouth, a region less informative about certain emotions, and certainly about where their interlocutor is looking (Neumann, Spezio, Piven & Adolphs, 2006). Further, they have impaired visual memory for faces (Blair, Frith, Smith, Abell & Cipolotti, 2002) and their neural responses to social stimuli do not show attentional modulation (Bird, Catmur, Silani, Frith & Frith, 2006). That being said, recent data have indicated that adults with Asperger's syndrome do not make initial saccades or show a looking bias towards the correct choice in a false belief task, even though they do not differ from normals in their performance in a verbally mediated theory of mind task (Senju, Southgate, White & Frith, 2009). In the same vein, it has been suggested that the correlation between verbal mental age and ability to pass a theory of mind task in children with autism may be a result of verbally mediated reasoning rather than social attunement (Happé, 1995). While these data may seem incongruent with the idea of an obligatory link between linguistic ability and proclivity towards social behavior, it is important to note that these are the results of compensatory mechanisms that come into play during or after wide-scale deficits in development. For instance, children who are able to use verbal reasoning to complete a theory of mind task form only a very small proportion of all autistic children and, further, often require a verbal mental age that is double that of normal children who are able to complete similar tasks (Happé, 1995). Thus, the fact that they are able to use some kind of cognitive compensation is indicative of the general functional plasticity of the brain, and is not at odds with the theory that the social motivation that these children seem to lack would have gated earlier (typical) maturation of verbal ability, as well as social reasoning.

Human infants with Williams's Syndrome have disrupted cortical development, including increased cortical thickness (Thompson, Lee, Dutton, Geaga, Hayashi, Eckert, Bellugi, Galaburda, Korenberg, Mills, Toga & Reiss, 2005) and mild to moderate mental retardation, but do not show linguistic disability of a magnitude that corresponds to these cortical anomalies (Bellugi, Lichtenberger, Jones, Lai & St George, 2000). As adults, these individuals show heightened sociability (Jones, Bellugi, Lai, Chiles, Reilly, Lincoln & Adolphs, 2000) and interest in faces (Laing, Butterworth, Ansari, Gsödl, Longhi, Panagiotaki, Paterson & Karmiloff-Smith, 2002) but don't seem to recruit amygdalar function in a manner akin to normals during face processing tasks (Paul, Snyder, Jaist, Raichle, Bellugi & Stiles, 2009). In general, there seems to be a distinct propensity for heightened affect, especially in social situations, such as those associated with linguistic narrative (Reilly, Losh, Bellugi & Wulfeck, 2004). That there is some abnormality is not being contended, and the argument here is *not* that individuals with Williams Syndrome have completely preserved linguistic skills in the face of extensive cortical aberrations that result in retardation, but that given the extensiveness of the cortical damage, the preservation of

even the compromised linguistic ability that these individuals display is surprising – and suggests the possible involvement of areas that lie outside the cortex in linguistic processes. Our suggestion is that spared or exaggerated social motivation in these individuals accounts for the *protection* of language.

We would like to note that both Autism spectrum disorders and William's Syndrome are complex and incompletely understood brain disorders and it would be beyond the scope of our current knowledge to draw definitive conclusions from the data we currently have, either in favor of social motivation as the driving force behind early language learning, or in favor of a disjunct between social and linguistic behavior. However, on the basis of emerging trends from current research, we would argue that studies of individuals with these disorders do seem to indicate that they provide us with populations where either the social system is not the *tour-de-force* that it normally is (Autism) or where hyperactivity of the social circuits is offsetting cognitive deficits in a manner that provides impetus to greater learning of faculties that are embedded within the social domain (William's Syndrome).

Finally, an aside about another genetic disorder associated with language deficiency – the alteration in the expression of FOXP2 producing serious deficiencies in the production of and structure of language in the KE family, as well as multiple deficiencies in praxis. While it is not yet known exactly what feature of brain organization its early FOXP2 expression supervises, the fact that it is under positive selection in both songbirds and humans, is expressed in the basal ganglia, and is associated with some aspect of fine temporal organization in both production and perception is interesting evidence of a further kind for 'developmental conservation' of paths of adaptation across phyla (Vargha-Khadem, Watkins, Alcock, Fletcher & Passingham, 1995; Enard, Przeworski, Fisher, Lai, Wiebe, Kitano, Monaco & Paabo, 2002; Haesler, Wada, Nshdejan, Morrisey, Lints, Jarvis & Scharff, 2004). This kind of important, facilitatory change for rapid motor production and perhaps perception points out the necessary caution that an argument highlighting the social, motivational aspects of communication systems does not in any way preclude other alterations making vocal (or other) communication more feasible, including the speed-of-processing aspect hinted at in the FOXP2 expression work, or the requirement for adequate memory stores in both birdsong and language.

What has changed in the human brain permitting language acquisition?

In this paper, we have attempted to draw together selected types of evidence about the organization and evolution of systems of social communication to seat human language learning in this context, ignoring the voluminous literature about the sensory, analytical, cognitive and linguistic

nature of language. We have reviewed evidence to suggest the elaborate interconnection of social and motivational circuitry, in neuroanatomical terms. Current evidence further suggests an unusual combinatorial capacity in motivational systems over evolutionary time (Goodson, 2005). We suggest that the new linkage that has occurred in the evolution of development has been made between the neural representation of central caregivers (as to a lamb, certain mate or to a zebra finch 'tutor'), the motivational systems, and cortical vocal and gestural learning systems. In species that have evolved vocal learning, influencing the desired individual's attention through behavior, particularly vocalization, as indicated by gaze, contact or localization produces the most profound reward. A learning system, cortical in origin, of unusual power, has forged a unique link to the social circuit whereby social motivation 'gates' vocal learning (Figure 3). Figure 3a presents a diagram of such a proposed 'gate' in the vole social behavior network whereby dopaminergic reward systems, through the Nucleus Accumbens, become associated with specific individuals – as in the case of monogamous voles – in concert with V1aR receptor upregulation in the Ventral Pallidum, while Figure 3b hypothesizes a similar link or gate between recognition of a parent and a vocal learning system.

A whole constellation of well-studied recognition systems stands in place to organize the recognition of those significant individuals who engage this newly connected, socially gated nervous system. Faces are special; a number of avenues suggest special evolution of both the morphology and the recognition of the significance of eye gaze and conformation (Senju & Johnson, 2009); recognition of the mother's voice occurs through learning *in utero* (DeCasper & Fifer, 1980); the preferences of infants for familiar tastes, smells and contact are well known.

Implications

The essential feature of the argument is that language emerges from a newly forged linkage of motivational systems and a desire to influence an object of much desire, in the presence of a powerful learning device. We have some evidence from early types of brain disorders that while near-catastrophic damage to sensory and cortical systems allows preservation of language, alterations to disposition or motivation are disproportionately damaging. For principally technical reasons at this point, direct evidence for disproportional activation of motivational systems during early language learning in infancy, such as that described in the activation of immediate early gene systems for the preferred partner in monogamous voles, cannot presently be gathered in human infants. Given the rate of technological advances in this area, we expect that this situation will not remain so for long, as a recent study using near infrared spectroscopy to examine the activation of the cortex in young

infants in social situations attests (Lloyd-Fox, Blasi, Volein, Everdell, Elwell & Johnson, 2009). In the near future we suggest that researchers in early language development turn their attention from the storage device, the cortex, to the basal forebrain and striatum, which provide the motivational structure for behavior.

A final cautionary note, which may in fact be a point of leverage for future research, is that the powerful social circuitry for particular behaviors we have described is rarely stable for the lifetime of the individual. Both parenthood and childhood pass, and the nature of those attachments mutates and special learning epochs disappear. In animal models, the neural signatures of those recognition, attachment and learning systems pass too. In fact, motivational coupling may entirely reverse, for example, when a formerly loved home and family become noxious to a newly mature animal seeking a new home turf. Learning a new vocabulary word in adolescence is rarely an orgasmic experience. Perhaps, however, we feel an echo of that earlier state when as adults we make ourselves understood in a new country, in a new language for the first time.

Acknowledgements

This work appears as part of the first author's dissertation to be submitted to the Faculty of the Graduate School of Cornell University. This research was supported by an NSF grant to the second author (NSF DBI 0848612). We would like to thank Jim Goodson, Morten Christiansen and Marcy Kingsbury for their careful critiques of, and helpful comments on, earlier versions of this manuscript.

References

- Adar, E., Lotem, A., & Barnea, A. (2008). The effect of social environment on singing behavior in the zebra finch (*Taeniopygia guttata*) and its implication for neuronal recruitment. *Behavioural Brain Research*, **187** (1), 178–184.
- Adret, P. (1993). Operant conditioning, song learning and imprinting to taped song in the zebra finch. *Animal Behaviour*, **46**, 149–159.
- Alcock, J. (2009). *Animal behavior: An evolutionary approach* (9th edn.). New York: Sinauer Associates.
- Bates, E.A. (2004). Explaining and interpreting deficits in language development across clinical groups: where do we go from here? *Brain and Language*, **88** (2), 248–253.
- Bates, E., Reilly, J., Wulfeck, B., Dronkers, N., Opie, M., Fenson, J., Kriz, S., Jeffries, R., Miller, L., & Herbst, K. (2001). Differential effects of unilateral lesions on language production in children and adults. *Brain and Language*, **79** (2), 223–265.
- Bellugi, U., Lichtenberger, L., Jones, W., Lai, Z., & St George, M.I. (2000). The neurocognitive profile of Williams syndrome: a complex pattern of strengths and weaknesses. *Journal of Cognitive Neuroscience*, **12** (Suppl. 1), 7–29.
- Biben, M., & Symmes, D. (1991). Playback studies of affiliative vocalizing in captive squirrel monkeys: familiarity as a cue to response. *Behaviour*, **117**, 1–19.
- Bigelow, A.E. (1998). Infants' sensitivity to familiar contingencies in social interaction. *Infant Behavior and Development*, **21**, 149–162.
- Bigelow, A.E., & Rochat, P. (2006). Two-month-old infants' sensitivity to social contingency in mother–infant and stranger–infant interaction. *Infancy*, **9** (3), 313–325.
- Bird, G., Catmur, C., Silani, G., Frith, C., & Frith, U. (2006). Attention does not modulate neural responses to social stimuli in autism spectrum disorders. *NeuroImage*, **31** (4), 1614–1624.
- Bowles, S. (2006). Group competition, reproductive leveling, and the evolution of human altruism. *Science*, **314** (5805), 1569–1572.
- Blair, R.J., Frith, U., Smith, N., Abell, F., & Cipolotti, L. (2002). Fractionation of visual memory: agency detection and its impairment in autism. *Neuropsychologia*, **40**, 108–118.
- Bloom, K., Russell, A., & Wessenberg, K. (1987). Turn taking affects the quality of infant vocalizations. *Journal of Child Language*, **14**, 211–227.
- Broad, K.D., Curley, J.P., & Keverne, E.B. (2006). Mother–infant bonding and the evolution of mammalian social relationships. *Philosophical Transactions of the Royal Society of London B*, **361**, 2199–2214.
- Cardinal, R.N., Parkinson, J.A., Hall, J., & Everitt, B.J. (2002). Emotion and motivation: the role of the amygdala, ventral striatum, and prefrontal cortex. *Neuroscience and Biobehavioral Reviews*, **26** (3), 321–352.
- Cheng, M.F. (2003). Vocal self-stimulation: from the ring dove story to emotion-based vocal communication. *Advances in the Study of Behavior*, **33**, 309–353.
- Cheng, M.F., & Durand, S.E. (2004). Song and the limbic brain: a new function for the bird's own song. *Annals of the New York Academy of Sciences*, **1016**, 611–627.
- Chomsky, N. (1981). *Rules and representations*. New York: Columbia University Press.
- Cohen, J., & Cheng, M.F. (1981). The role of the midbrain in courtship behavior of the female ring dove (*Streptopelia risoria*): evidence from radiofrequency lesion and hormone implant studies. *Brain Research*, **207**, 279–301.
- Cohen, M.X., Schoene-Bake, J.C., Elger, C.E., & Weber, B. (2009). Connectivity-based segregation of the human striatum predicts personality characteristics. *Nature Neuroscience*, **12** (1), 32–34.
- Conboy, B.T., Brooks, R., Meltzoff, A.N., & Kuhl, P.K. (2008). Joint engagement with language tutors predicts brain and behavioral responses to second-language phonetic stimuli. Paper presented at the XVIth International Conference on Infant Studies, Vancouver.
- Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. In J.H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 163–228). New York: Oxford University Press.
- Damasio, H., Grabowski, T., Frank, R., Galaburda, A.M., & Damasio, A.R. (1994). The return of Phineas Gage: clues about the brain from the skull of a famous patient. *Science*, **264** (5162), 1102–1105.
- DeCasper, A.J., & Fifer, W.P. (1980). Of human bonding: newborns prefer their mothers' voices. *Science*, **208** (4448), 1174–1176.

- Depue, R.A., & Morrone-Strupinsky, J.V. (2005). A neurobehavioral model of affiliative bonding: implications for conceptualizing a human trait of affiliation. *Behavioral and Brain Sciences*, **28**, 313–350.
- Doupe, A.J., & Kuhl, P.K. (1999). Birdsong and human speech: common themes and mechanisms. *Annual Review of Neuroscience*, **22**, 567–631.
- Dunbar, R.I.M., & Shultz, S. (2007). Understanding primate brain evolution. *Philosophical Transactions of the Royal Society of London, B*, **362**, 649–658.
- Eales, L.A. (1989). The influence of visual and vocal interaction on song learning in zebra finches. *Animal Behaviour*, **37**, 507–508.
- Edinger, L. (1888–1903). *Investigations on the comparative anatomy of the brain* (Vols 1–5). (Translation from German.) Frankfurt/Main: Moritz Diesterweg.
- Elman, J.L., Bates, E.A., & Johnson, M.H. (1996). *Rethinking innateness: A connectionist perspective on development*. Cambridge, MA: MIT Press.
- Enard, W., Przeworski, M., Fisher, S.E., Lai, C.S.L., Wiebe, V., Kitano, T., Monaco, A.P., & Paabo, S. (2002). Molecular evolution of FOXP2, a gene involved in speech and language. *Nature*, **418**, 869–872.
- Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development*, **8** (2), 181–195.
- Fernald, A. (1992a). Human maternal vocalisations to infants as biologically relevant signals: an evolutionary perspective. In J.H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 391–428). New York: Oxford University Press.
- Fernald, A. (1992b). Meaningful melodies in mothers' speech to infants. In H. Papousek, U. Jurgens, & M. Papousek (Eds.), *Nonverbal vocal communication: Comparative and developmental approaches* (pp. 262–282). Cambridge: Cambridge University Press.
- Fernald, A., Kermanschachi, N., & Lees, D. (1984). The rhythms and sounds of soothing: maternal vestibular, tactile, and auditory stimulation and infant state. International Conference on Infant Studies, New York.
- Fernald, A., & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology*, **20**, 106–112.
- Fernald, A., Taeschner, T., Dun, J., Papousek, M., de Boysson-Bardies, B., & Fukui, I. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language*, **16**, 477–501.
- Fodor, J.A. (1983). *The modularity of mind: An essay on faculty psychology*. Cambridge, MA: MIT Press.
- Fromkin, V., Krashen, S., Curtis, S., Rigler, D., & Rigler, M. (1974). The development of language in Genie: a case of language acquisition beyond the 'critical period'. *Brain and Language*, **1**, 81–107.
- Goldman-Rakic, P.S. (1999). The physiological approach: functional architecture of working memory and disordered cognition in schizophrenia. *Biological Psychiatry*, **46** (5), 650–661.
- Goldstein, M.H., King, A.P., & West, M.J. (2003). Social interaction shapes babbling: testing parallels between bird-song and speech. *Proceedings of the National Academy of Sciences, USA*, **100** (13), 8030–8035.
- Goldstein, M.H., & Schwade, J.A. (2008). Social feedback to infants' babbling facilitates rapid phonological learning. *Psychological Science*, **19** (5), 515–523.
- Goodson, J.L. (2005). The vertebrate social behavior network: evolutionary themes and variations. *Hormones and Behavior*, **48**, 11–22.
- Goodson, J.L., & Evans, A.K. (2004). Neural responses to territorial challenge and nonsocial stress in male song sparrows: segregation, integration, and modulation by a vasopressin V1 antagonist. *Hormones and Behavior*, **46**, 371–381.
- Goodson, J.L., Evans, A.K., Lindberg, L., & Allen, C.D. (2005). Neuro-evolutionary patterning of sociality. *Proceedings of the Royal Society of London, B*, **272**, 227–235.
- Haesler, S., Wada, K., Nshdejan, A., Morrisey, E.E., Lints, T., Jarvis, E.D., & Scharff, C. (2004). FoxP2 expression in avian vocal learners and non-learners. *Journal of Neuroscience*, **24**, 3164–3175.
- Hammock, E.A.D., & Young, L.J. (2004). Functional microsatellite polymorphism associated with divergent social structure in vole species. *Molecular Biology and Evolution*, **21** (6), 1057–1063.
- Hammock, E.A.D., & Young, L.J. (2005). Microsatellite instability generates diversity in brain and sociobehavioral traits. *Science*, **308** (5728), 1630–1634.
- Happe, F.G. (1995). The role of age and verbal ability in the theory of mind task performance of subjects with autism. *Child Development*, **66** (3), 843–855.
- Hara, E., Kubikova, L., Hessler, N.A., & Jarvis, E.D. (2007). Role of midbrain dopaminergic system in the modulation of vocal brain activation by social context. *European Journal of Neuroscience*, **25**, 3406–3416.
- Hausberger, M., Richard-Yris, M., Henry, L., Lepage, L., & Schmidt, I. (1995). Song sharing reflects the social organization in a captive group of European starlings (*Sturnus vulgaris*). *Journal of Comparative Psychology*, **109**, 222–241.
- Hessler, N.A., & Doupe, A.J. (1999). Social context modulates singing-related neural activity in the songbird forebrain. *Nature Neuroscience*, **2**, 209–211.
- Houx, B.B., & ten Cate, C. (1999). Learning from playback tutor song in zebra finches: is there an effect of operant contingency? *Animal Behaviour*, **57**, 837–845.
- Huang, Y.C., & Hessler, N.A. (2008). Social modulation during songbird courtship potentiates midbrain dopaminergic neurons. *PLoS One*, **3** (10), e3281.
- Immelmann, K. (1969). Song development in the zebra finch and other estrildid finches. In R.A. Hinde (Ed.), *Bird vocalizations* (pp. 61–74). London: Cambridge University Press.
- Insel, T.R., Wang, Z., & Ferris, C.F. (1994). Patterns of brain vasopressin receptor distribution associated with social organization in microtine rodents. *Journal of Neuroscience*, **14**, 5381–5392.
- Jarvis, E.D. (2007). The evolution of vocal learning systems in birds and humans. In Jon H. Kaas (Ed.), *Evolution of nervous systems* (pp. 213–228). New York: Academic Press.
- Jarvis, E.D., Gunturkun, O., Bruce, L., Csillag, A., Karten, H., Kuenzel, W., et al. (2005). Avian brains and a new understanding of vertebrate brain evolution. *Nature Reviews Neuroscience*, **6** (2), 151–159.
- Johnston, R.E. (2005). Communication by mosaic signals: individual recognition and underlying neural mechanisms. In R. Mason (Ed.), *Chemical signals in vertebrates* (Vol. 10, pp. 269–282). New York: Kluwer/Plenum Press.
- Jones, W., Bellugi, U., Lai, Z., Chiles, M., Reilly, J., Lincoln, A., & Adolphs, R. (2000). II. Hypersociability in Williams syndrome. *Journal of Cognitive Neuroscience*, **12** (Suppl. 1), 30–46.

- Karten, H.J. (1997). Evolutionary developmental biology meets the brain: the origins of the mammalian cortex. *Proceedings of the National Academy of Sciences, USA*, **94**, 2800–2804.
- Kirschner, M.W., & Gerhart, J.C. (2005). *The plausibility of life: Resolving Darwin's dilemma*. New Haven, CT: Yale University Press.
- Klinger, L., Dawson, G., & Renner, P. (2002). Autistic disorder. In E. Mash & R. Barkley (Eds.), *Child psychopathology* (2nd edn., pp. 409–454). New York: Guilford Press.
- Kuhl, P.K. (2007). Is speech learning 'gated' by the social brain? *Developmental Science*, **10**, 110–120.
- Kuhl, P.K., Coffey-Corina, S., Padden, D., & Dawson, G. (2005). Links between social and linguistic processing of speech in preschool children with autism: behavioral and electrophysiological evidence. *Developmental Science*, **8**, F1–F12.
- Kuhl, P.K., Tsao, F.M., & Liu, H.M. (2003). Foreign-language experience in infancy: effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences, USA*, **100**, 9096–9101.
- Laing, E., Butterworth, G., Ansari, D., Gsödl, M., Longhi, E., Panagiotaki, G., Paterson, S., & Karmiloff-Smith, A. (2002). Atypical development of language and social communication in toddlers with Williams syndrome. *Developmental Science*, **5** (2), 233–246.
- Lane, H.L. (1976). *The wild boy of Aveyron*. Cambridge, MA: Harvard University Press.
- Lebreton, M., Barnes, A., Miettunen, J., Peltonen, L., Ridler, K., Veijola, J., et al. (2009). The brain structural disposition to social interaction. *European Journal of Neuroscience*, **29**, 2247–2252.
- Lim, M.M., Hammock, E.A.D., & Young, L.J. (2004a). The role of vasopressin in the genetic and neural regulation of monogamy. *Journal of Neuroendocrinology*, **16** (4), 325–332.
- Lim, M.M., Murphy, A.Z., & Young, L.J. (2004b). Ventral striato-pallidal oxytocin and vasopressin V1a receptors in the monogamous prairie vole (*Microtus ochrogaster*). *Journal of Comparative Neurology*, **468**, 555–570.
- Lim, M.M., & Young, L.J. (2002). Blockade of vasopressin V1a receptors in the ventral pallidum prevents partner preference formation in monogamous prairie voles. *Society for Neuroscience Abstract*, **28**, 89.2. Orlando, FL: 32nd Annual Meeting.
- Lim, M.M., & Young, L.J. (2004). Vasopressin-dependent neural circuits underlying pair bond formation in the monogamous prairie vole. *Neuroscience*, **125**, 35–45.
- Lipkind, D., Nottebohm, F., Rado, R., & Barnea, A. (2002). Social change affects the survival of new neurons in the forebrain of adult songbirds. *Behavioral Brain Research*, **133**, 31–43.
- Lloyd-Fox, S., Blasi, A., Volein, A., Everdell, N., Elwell, C.E., & Johnson, M.H. (2009). Social perception in infancy: a near infrared spectroscopy study. *Child Development*, **80** (4), 986–999.
- Locke, J. (1993). *The path to spoken language*. Cambridge, MA: Harvard University Press.
- Mann, N.I., & Slater, P.J.B. (1995). Song tutor choice by zebra finches in aviaries. *Animal Behaviour*, **49** (3), 811–820.
- Marler, P. (1970). A comparative approach to vocal learning: song development in white-crowned sparrows. *Journal of Comparative Physiology and Psychology*, **71**, 1–25.
- Marler, P., & Sherman, V. (1985). Innate differences in singing behavior of sparrows reared in isolation from adult conspecific song. *Animal Behaviour*, **33**, 57–71.
- Martinez-Garcia, F., Martinez-Marcos, A., & Lanuza, E. (2002). The pallial amygdala of amniote vertebrates: evolution of the concept, evolution of the structure. *Brain Research Bulletin*, **57** (3–4), 463–469.
- Martinez-Garcia, F., Novejarque, A., & Lanuza, E. (2007). Evolution of the amygdala in vertebrates. In Jon H. Kaas (Ed.), *Evolution of nervous systems* (pp. 255–334). New York: Academic Press.
- Medina, L. (2007). Do birds and reptiles possess homologues of mammalian visual, somatosensory and motor cortices? In Jon H. Kaas (Ed.), *Evolution of nervous systems* (Vol. 2, pp. 163–194). New York: Academic Press.
- Mogenson, G.J., Jones, D.L., & Yim, C.Y. (1980). From motivation to action: functional interface between the limbic system and the motor system. *Progress in Neurobiology*, **14** (2–3), 69–97.
- Moll, H., & Tomasello, M. (2007). Cooperation and human cognition: the Vygotskian intelligence hypothesis. *Philosophical Transactions of the Royal Society of London, B*, **362** (1480), 639–648.
- Morgane, P.J., Galler, J.R., & Mokler, D.J. (2005). A review of systems and networks of the limbic forebrain/limbic mid-brain. *Progress in Neurobiology*, **75**, 143–160.
- Nacewicz, B.M., Dalton, K.M., Johnstone, T., Long, M.T., McAuliff, E.M., Oakes, T.R., Alexander, A.L., & Davidson, R.J. (2006). Amygdala volume and nonverbal social impairment in adolescent and adult males with autism. *Archives of General Psychiatry*, **63**, 1417–1428.
- Neumann, D., Spezio, M.L., Piven, J., & Adolphs, R. (2006). Looking you in the mouth: abnormal gaze in autism resulting from impaired top-down modulation of visual attention. *Social Cognitive and Affective Neuroscience*, **1** (3), 194–202.
- Newman, S.W. (1999). The medial extended amygdala in male reproductive behavior: a node in the mammalian social behavior network. *Annals of the New York Academy of Sciences*, **877**, 242–257.
- Nowicki, S. (1989). Vocal plasticity in captive black-capped chickadees: the acoustic basis of call convergence. *Animal Behaviour*, **37**, 64–73.
- Osterling, J., & Dawson, G. (1994). Early recognition of children with autism: a study of first birthday home videotapes. *Journal of Autism and Developmental Disorders*, **24**, 247–257.
- Osterling, J., Dawson, G., & Munson, J. (2002). Early recognition of 1-year-old infants with autism spectrum disorder versus mental retardation. *Development and Psychopathology*, **14**, 239–251.
- Owren, M.J., Dieter, J.A., Seyfarth, R.M., & Cheney, D.L. (1993). Vocalizations of rhesus (*Macaca mulatta*) and Japanese (*M. fuscata*) macaques cross-fostered between species show evidence of only limited modification. *Developmental Psychobiology*, **26**, 389–406.
- Pagel, M. (2009). Human language as a culturally transmitted replicator. *Nature Reviews Genetics*, **10** (6), 405–415.
- Papousek, M., Papousek, H., & Bornstein, M.H. (1985). The naturalistic vocal environment of young infants: on the significance of homogeneity and variability in parental speech. In T. Field & N. Fox (Eds.), *Social perception in infants* (pp. 269–297). Norwood, NJ: Ablex.
- Paul, B.M., Snyder, A.Z., Haist, F., Raichle, M.E., Bellugi, U., & Stiles, J. (2009). Amygdala response to faces parallels social behavior in Williams syndrome. *Social Cognitive and Affective Neuroscience*, **4** (3), 278–285.

- Pepperberg, I.M. (1999). *The Alex studies: Cognitive and communicative abilities of grey parrots*. Cambridge, MA: Harvard University Press.
- Preston, S.D., & de Waal, F.B.M. (2001). Empathy: its ultimate and proximate bases. *Behavioral and Brain Sciences*, **25** (1), 1–20.
- Puelles, L. (2001). Thoughts on the development, structure and evolution of the mammalian and avian telencephalic pallium. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences*, **356** (1414), 1583–1598.
- Puelles, L., Kuwana, E., Puelles, E., Bulfone, A., Shimamura, K., Keleher, J., Smiga, S., & Rubenstein, J.L.R. (2000). Pallial and subpallial derivatives in the embryonic chick and mouse telencephalon, traced by the expression of the genes *Dlx-2*, *Emx-1*, *Nkx-2.1*, *Pax-6*, and *Tbr-1*. *Journal of Comparative Neurology*, **424**, 409–438.
- Reilly, J., Losh, M., Bellugi, U., & Wulfeck, B. (2004). 'Frog, where are you?' Narratives in children with specific language impairment, early focal brain injury and Williams syndrome. *Brain and Language*, **88**, 229–247.
- Reiner, A.J. (2000). A hypothesis as to the organization of cerebral cortex in the common amniote ancestor of modern reptiles and mammals. in *Evolutionary developmental biology of the cerebral cortex. Novartis Foundation Symposium 228* (Eds G.R. Bock & G. Cardew) (pp. 83–108). Chichester: Wiley.
- Richerson, P.J., & Boyd, R. (2005). *Not by genes alone: How culture transformed human evolution*. Chicago, IL: University of Chicago Press.
- Roush, R.S., & Snowdon, C.T. (1994). Ontogeny of food-associated calls in cotton-top tamarins. *Animal Behaviour*, **47**, 263–273.
- Roush, R.S., & Snowdon, C.T. (1999). The effects of social status on food-associated calling behaviour in captive cotton-top tamarins. *Animal Behaviour*, **58** (6), 1299–1305.
- Sasaki, A., Sotnikova, T.D., Gainetdinov, R.R., & Jarvis, E.D. (2006). Social context-dependent singing-regulated dopamine. *Journal of Neuroscience*, **26**, 9010–9014.
- Schumann, C.M., Hamstra, J., Goodlin-Jones, B.L., Lotspeich, L.J., Kwon, H., Buonocore, M.H., Lammers, C.R., Reiss, A.L., & Amaral, D.G. (2004). The amygdala is enlarged in children but not adolescents with autism; the hippocampus is enlarged at all ages. *Journal of Neuroscience*, **24**, 6392–6401.
- Senju, A., & Johnson, M.H. (2009). The eye contact effect: mechanisms and development. *Trends in Cognitive Sciences*, **13** (3), 127–134.
- Senju, A., Southgate, V., White, S., & Frith, U. (2009). Mindblind eyes: an absence of spontaneous theory of mind in Asperger syndrome. *Science*, **325**, 883–885.
- Seyfarth, R.M., & Cheney, D.L. (1986). Vocal development in vervet monkeys. *Animal Behaviour*, **34**, 1640–1658.
- Seyfarth, R.M., & Cheney, D.L. (1997). Behavioral mechanisms underlying vocal communication in nonhuman primates. *Animal Learning & Behavior*, **25**, 249–267.
- Smith, H.J., Newman, J.D., Hoffman, H.J., & Fetterley, K. (1982). Statistical discrimination among vocalizations of individual squirrel monkeys (*Saimiri sciureus*). *Folia Primatologica*, **37**, 267–279.
- Smolker, R.A., & Pepper, J.W. (1999). Whistle convergence among allied male bottlenose dolphins (Delphinidae, Tursiops sp.). *Ethology*, **105**, 595–617.
- Snowdon, C.T., & Elowson, A.M. (1994). Pygmy Marmosets, *Cebuella pygmaea*, modify vocal structure in response to changed social environment. *Animal Behaviour*, **47**, 1267–1277.
- Snowdon, C.T., Elowson, A.M., & Roush, R.S. (1997). Social influences on vocal development in New World Primates. In C.T. Snowdon & M. Hausberger (Eds.), *Social influences on vocal development* (pp. 234–248). Cambridge: Cambridge University Press.
- Stern, D.N., Spieker, S., & MacKain, K. (1982). Intonation contours as signals in maternal speech to prelinguistic infants. *Developmental Psychology*, **18** (5), 727–749.
- Striedter, G.F. (1997). The telencephalon of tetrapods in evolution. *Brain, Behavior and Evolution*, **49** (4), 179–213.
- Thompson, P.M., Lee, A.D., Dutton, R.A., Geaga, J.A., Hayashi, K.M., Eckert, M.A., Bellugi, U., Galaburda, A.M., Korenberg, J.R., Mills, D.L., Toga, A.W., & Reiss, A.L. (2005). Abnormal cortical complexity and thickness profiles mapped in Williams syndrome. *Journal of Neuroscience*, **25** (16), 4146–4158.
- Thorpe, W.H. (1958). The learning of song patterns by birds with especial reference to the song of the chaffinch, *Fringilla coelebs*. *Ibis*, **100**, 535–570.
- Tomasello, M. (1999). *The cultural origins of human cognition*. Cambridge, MA: Harvard University Press.
- Vargha-Khadem, F., Watkins, K., Alcock, K., Fletcher, P., & Passingham, R. (1995). Praxic and nonverbal cognitive deficits in a large family with a genetically transmitted speech and language disorder. *Proceedings of the National Academy of Sciences of the United States of America*, **92**, 930–933.
- Wang, Z., Zhou, L., Hulihian, T.J., & Insel, T.R. (1996). Immunoreactivity of central vasopressin and oxytocin pathways in microtine rodents: a quantitative comparative study. *Journal of Comparative Neurology*, **366**, 726–737.
- Warneken, F., & Tomasello, M. (2008). The roots of human altruism. *British Journal of Psychology*, **21**, 1–18.
- Warneken, F., & Tomasello, M. (2009). Varieties of altruism in children and chimpanzees. *Trends in Cognitive Sciences*, **13**, 397–402.
- West, M.J., & King, A.P. (1988). Female visual displays affect the development of male song in the cowbird. *Nature*, **334** (6179), 244–246.
- Wilson, D.S., Van Vugt, M., & O'Gorman, R. (2008). Multi-level selection theory and major evolutionary transitions: implications for psychological science. *Current Directions in Psychological Science*, **17** (1), 6–9.
- Winter, P., Handley, P., Ploog, D., & Schott, D. (1973). Ontogeny of squirrel monkey calls under normal conditions and under acoustic isolation. *Behaviour*, **47**, 230–239.
- Wong Fillmore, L. (1991). When learning a second language means losing the first. *Early Childhood Research Quarterly*, **6**, 323–346.
- Young, L.J., Nilsen, R., Waymire, K.G., MacGregor, G.R., & Insel, T.R. (1999). Increased affiliative response to vasopressin in mice expressing the V1a receptor from a monogamous vole. *Nature*, **400**, 766–768.
- Young, L.J., & Wang, Z. (2004). The neurobiology of pair bonding. *Nature Neuroscience*, **7** (10), 1048–1054.

Received: 6 August 2009

Accepted: 10 May 2010