



Mirror neurons, birdsong, and human language: a hypothesis

Florence Levy *

School of Psychiatry, University of New South Wales and Prince of Wales Hospital, Sydney, NSW, Australia

Edited by:

Alan Apter, Schneider Children's
Medical Center of Israel, Israel

Reviewed by:

Jens Benninghoff, Ludwig
Maximilians University, Germany
Marco Grados, Johns Hopkins
University School of Medicine, USA

***Correspondence:**

Florence Levy, Head, Child and
Family East, Prince of Wales Hospital,
Sydney, NSW 2025, Australia.
e-mail: f.levy@unsw.edu.au

The mirror system hypothesis and investigations of birdsong are reviewed in relation to the significance for the development of human symbolic and language capacity, in terms of three fundamental forms of cognitive reference: iconic, indexical, and symbolic. Mirror systems are initially iconic but can progress to indexical reference when produced without the need for concurrent stimuli. Developmental stages in birdsong are also explored with reference to juvenile subsong vs complex stereotyped adult syllables, as an analogy with human language development. While birdsong remains at an indexical reference stage, human language benefits from the capacity for symbolic reference. During a pre-linguistic “babbling” stage, recognition of native phonemic categories is established, allowing further development of subsequent prefrontal and linguistic circuits for sequential language capacity.

Keywords: mirror neurons, birdsong, symbolic reference, language

There has been considerable interest in the functions of visual-motor neurons that fire when an animal performs a goal-directed action. The present review explores mirror system and birdsong developments as an analogy to stages of human prefrontal and language development. While the development of prefrontal cortex is suggested as being important for working memory and cognitive development, an understanding of mirror system and birdsong development may provide insights into human language development.

Deacon (1997) has described three fundamental forms of cognitive reference: iconic, indexical, and symbolic. He points out, that landscapes, portraits of all kinds are iconic of what they depict. However when we say something is an “index,” we mean that it is somehow causally linked to something else. Deacon points out that most forms of animal communication are indexical, from pheromonal odors (that indicate an animal’s physiological state or proximity), to alarm calls (that indicate the presence of a dangerous predator). On the other hand, a symbol is derived from “social convention, tacit agreement, or explicit code, which establishes the relationship that links one thing to another” (Deacon, 1997).

Based on the work of Peirce (1931–1958), Deacon points out that the difference between the three different modes of reference can be understood in terms of levels of interpretation, arranged in an ascending order that reflects a prior competence. “In other words, reference itself is hierarchic in structure; more complex forms of reference are built up from simpler forms.” Thus “indexical reference depends upon iconic reference and symbolic reference depends upon indexical reference” (Deacon, 1997).

Deacon (1997) suggests that a shift from associative (indexical) to symbolic predictions is initially a change in mnemonic strategy that results in a radical transformation in the mode of representation. “What one knows in one way, gets recoded in another way. It gets re-represented. . . And because this recoding is based on higher-order relationships, not individual details. It often vastly

simplifies the mnemonic problem and vastly augments the representational possibilities. . . Because the combinatorial rules encode not objects, but ways in which objects can be related, new symbols can immediately become incorporated and combined with others. The system of representational relationships between symbols as symbol systems grow, comprise an ever more complex matrix” (Deacon, 1997).

The above classification is useful in examining the relationship between mirror neurons, birdsong, and human prefrontal development and in the understanding of symbolic reference in human language.

MIRROR SYSTEM HYPOTHESIS

The discovery of mirror neurons that fire when an animal performs a goal-directed action or sees others perform the action gave rise to both monkey and human studies on the neural correlates of imitation (Rizzolatti and Craighero, 2004; Iacoboni and Dapretto, 2006). While the circuitry for motor imitation is described as encompassing the posterior inferior gyrus (including area F5), and adjacent ventral premotor cortex, as well as a posterior area located in the rostral part of the inferior parietal lobule, mirror neurons are also thought to have importance in language development. According to Iacoboni and Dapretto (2006) this evolutionary argument is based on the homology between area F5 of the macaque brain and Brodmann area 44 in the posterior inferior frontal gyrus of the human brain, an area linked with language.

Iacoboni and Dapretto (2006) also point out while trans-cranial magnetic stimulation (TMS) has demonstrated bilaterality for the human mirror system (MNS), the question is raised how a relatively bilateral system for action observation and imitation contributes to a left-lateralized system for language. The authors describe a further TMS study which investigated motor facilitation in the two hemispheres, while listening to acoustic sounds. The authors point out that as mirror neurons also respond to the

sound of an action, the motor system of listeners should be facilitated when listening to the sounds produced by actions, but such facilitation only occurred in the left hemisphere. This suggested that the left hemisphere of the human brain has a multimodal (visual, auditory) MNS, whereas the right hemisphere has only a visual MNS. "In humans, the shift from a purely visual to a multimodal MNS, could have determined functional changes that could have facilitated language and a left-lateralization of language functions" (Peirce, 1931–1958).

Arbib (2010) has proposed a mirror system hypothesis (MSH) of language evolution. He suggests that the mechanisms which support language in the human evolved atop a basic mechanism not originally related to communication, namely the mirror system for grasping. In the macaque monkey brain an area F5 just in front of the primary motor cortex (thought to be homologous with F1 in humans) contains neurons, active during manual and orofacial actions. According to Arbib, a subset of these neurons (mirror neurons), are also active when the monkey observes actions such as a precision pinch or a power grasp by a monkey or a human. Similarly, a "mirror system for grasping" has been shown in the human brain. Most significantly frontal activation was found in or near Broca's area, a region which in most humans lies in the left hemisphere and traditionally is associated with speech production. Arbib hypothesized seven stages in the evolution of languages, with the first three stages S1 (grasping), S2 (grasping shared with the common ancestor of human and monkey), S3 (a simple imitation system for grasping shared with the common ancestor) thought to be pre-hominid, and the next three S4 (a complex imitation system for grasping), S5 (protosign, a manual-based communication system, breaking through the fixed repertoire of primate vocalizations to yield an open repertoire), S6 (protospeech, the ability of control mechanisms evolved for protosign coming to control the vocal apparatus with increasing flexibility, thought to distinguish the hominid line from that of the great apes), while the final stage S7 is that of language (Arbib and Rizzolatti, 1997). According to Arbib, MSH is simply the assertion that the mechanisms which get us to the role of Broca's area in language depend in a crucial way on the mechanisms established in Stage 2, namely mirror mechanisms.

While the mirror system for grasping is described as evolving to support protosign and protospeech in humans (Arbib and Rizzolatti, 1997; Rizzolatti and Arbib, 1998; Iacoboni and Dapretto, 2006; Arbib, 2010). Arbib and Bota (2010) distinguish between the neural representation of the "sign" (as distinct from symbol), which inherits mirror properties linking the production of vocal, manual, and/or facial gestures for a word, on one hand and the phonological loop and working memory systems on the other (Baddeley, 2003). Arbib and Bota suggest that the development of grammar involves the notion that "Broca's area must be linked into prefrontal cortical (PFC) planning (and its administration by the basal ganglia) to assemble verb-argument and more complex hierarchical structures, finding the words, and binding them correctly." By implication, the "mirror" aspect of language suggests an early feed-forward "pantomime" stage, while the development of grammar implies a sequential recursive capacity based on PFC developments (Arbib and Bota, 2010).

Fitch (2005) has pointed out that the "weakest link" in the Arbib MNS model is the crucial link from protosign to protospeech, "specifically his elision between two distinct forms of imitation: vocal and manual." Fitch suggests that the co-evolution of vocal and manual gesture may have been more closely tied to music and dance than pantomime and linguistic communication.

"By this hypothesis, the crucial first step in human evolution from our last common ancestor with chimpanzees was the development of vocal imitation, similar in form and function to that independently evolved in many other vertebrate lineages (including cetaceans, pinnipeds, and multiple avian lineages)... This hypothetical musical protolanguage preceded any truly linguistic system, capable of communicating particulate propositional meanings... while dogs, birds, and apes can learn to map between meanings and words presented in isolation, the ability to extract words from arbitrary complex contexts and to recombine them in equally complex novel contexts is unattested in any non-human animal" (Fitch, 2005). Fitch points out that each generation of human children makes this "analytic" leap by the age of three, without tutelage, feedback, or specific scaffolding, in contrast with skills such as alphabetic writing.

In terms of Deacon's (1997) classification above, mirror mechanisms should originally be classified as iconic, but according to Arbib and Bota (2010), they provide a platform on which symbolic language representation may be built. Where mirror systems reproduce a motor or vocal action in the absence of a concurrent stimulus, they have progressed to an indexical representation.

BIRDSONG ANALOGY

A further analogy with early language development is provided by the investigation of birdsong.

For example, Aronov et al. (2008) point out that babbling is an early behavior produced by juveniles of vocal mammals and birds. While much of the brain of birds from spinal cord to midbrain reflects an organization common to most vertebrates, the "higher" brain regions including the forebrain are different. However, there are large nuclear masses, which resemble the mammalian basal ganglia (striatum), and a laminated isocortex (gray matter), which is separated from the underlying basal ganglia by a band of myelinated axons (white matter). Like humans, songbirds are dependant on hearing early in life for successful vocal learning. Birdsong and language both consist of ordered strings of sounds, separated by brief silent intervals. Song syllables are usually grouped together to form phrases or motifs (Aronov et al., 2008).

In zebra finches, babbling (called subsong) occurs roughly from 30 to 45 days post-hatch (dph). "Plastic-song" follows, with the gradual appearance of distinctive identifiable, but variable, vocal elements (syllables). According to the authors, plastic-song is by 90 dph gradually transformed into highly complex, stereotyped, motifs, or sequences of syllables that constitute adult song. The premotor circuit for adult song production is believed to consist of the high vocal center (HVC), robust nucleus of the archipallium (RA), and brainstem motor nuclei. This "motor pathway" is crucial for generating stereotyped, learned vocalizations, and exhibits firing, that is precisely time-locked to the song output (Aronov et al., 2008).

According to Aronov et al. (2008) another circuit, the anterior forebrain pathway (AFP) is homologous to the basal ganglia–thalamo-cortical loops in mammals, and projects to RA through a forebrain nucleus, lateral magnocellular nucleus of the nidopallium (LMAN). “Although LMAN is not required for singing in adult birds, it is necessary for normal song learning in juveniles, and plays a role in producing song variability in adult and juvenile birds” (Aronov et al., 2008). Aronov and colleagues eliminated the HVC bilaterally (important in adult singing) in nine subsong (33–44 dph), and in three additional birds in which they left the HVC intact, but specifically eliminated its projection to RA. After these manipulations, all young birds continued producing largely unaffected subsong. Also 12 older birds in the plastic-song stage (45–73 dph) and five adults also sang after HVC elimination, but lost structure and stereotypy and reverted to subsong-like vocalizations. In addition, when HVC was pharmacologically inactivated this reversion was fast and reversible, suggesting an immediate rather than long-term circuit change. The investigators posited three possibilities in relation to subsong: it is entirely produced by the midbrain or brainstem; it is driven by circuitry intrinsic to RA, even in the absence of HVC and LMAN; and third it is driven by or requires inputs from LMAN or RA. They tested these hypotheses by lesions or inactivations of RA and LMAN. The investigators found that RA lesions entirely blocked singing in juvenile birds ($n = 5$, 35–73 dph), indicating that subsong-like vocalizations required descending inputs from forebrain. Similarly, song production was abolished by lesions of the HVC and subsequent inactivation of LMAN ($n = 5$, 51–75 dph), indicating that RA circuitry without its afferent paths was not sufficient to generate singing (Aronov et al., 2008).

The above authors concluded that LMAN and possibly other components of the AFP constitute an essential premotor circuit for the production of early “babbling.” At the same time, the classical premotor nucleus HVC was not necessary for the generation of subsong. They proposed two premotor pathways in the songbird function, to produce vocalizations at different stages of development. “In young juveniles, the AFP generates poorly structured subsong, whereas in adult birds, the classical HVC-motor pathway generates highly stereotypic motor sequences. These pathways interact in the intermediate song stage to generate structured but variable vocalizations, upon which vocal learning operates.” The transfer of functional dominance from one pathway to another during vocal learning elegantly parallels their anatomical development. HVC does not reach its adult size until the late plastic-song stage; and establishes synapses in RA later than LMAN does “Song maturation and the decrease in vocal variability have thus been attributed to the strengthening of inputs from HVC and the concurrent weakening of inputs from LMAN.” The authors suggest that rather than a “neuronal group selection theory” of development (in which early motor behaviors originate in the same circuits that later produce adult behavior), their findings suggest that distinct specialized circuits are dedicated to production of highly variable juvenile behavior. That is, juvenile singing is driven by a circuit, *distinct* from that which produces adult behavior (Aronov et al., 2008).

The hypothesis by Aronov et al. (2008) that distinct cortical/subcortical circuits for the production of infant behavior

may be a general feature of developmental learning in the vertebrate brain, is important for the present review. The childhood to adolescent development of cortical/subcortical behavioral circuits involved in infant behaviors such as babbling, free play, and “over-activity,” with subsequent transition to goal-directed behavior is fundamental to the present concept of motor and language development. Like song maturation, the mechanisms by which this development is established may involve a pre-linguistic “babbling” stage, with subsequent transition to “goal-directed” language (Doupe and Kuhl, 1999).

Prather et al. (2008) describe two distinct populations of projection neurons in the swamp sparrow’s telencephalic nucleus (HVC), necessary for singing and normal song perception. These consist of HVC_{RA} cells which innervate song premotor neurons in the robust nucleus of the arcopallium (RA), and another HVC_X, that innervates a striatal region of the avian basal ganglia (area X⁶), important to song learning and perception.

The investigators looked at whether the activity in the HVC_X cells during singing was due to auditory feedback or a corollary of the song motor activity. It was noted that a period of playback “overlap” was locked precisely to features of the syllable being sung, suggesting corollary motor activity. Additionally, almost all HVC_X cells responded to only naturally occurring sequences, indicating that a sequence of at least two notes was necessary to elicit an auditory response. This selective auditory responsiveness of HVC_X cells extended to similar vocal sequences produced by other birds, making auditory–vocal HVC_X neurons well-suited to a role in communication (Prather et al., 2008).

“In many regards, auditory-vocal HVC_X cells are similar to visual-motor neurons in the monkey frontal cortex that are hypothesized to play a role in perception of human gestures, including human speech. In that light, the precise temporal alignment of auditory and vocal activity in HVC_X cells suggests that auditory-vocal mirror neurons express an additional mode of sensory-motor correspondence not previously reported for visual-motor mirror neurons” (Prather et al., 2008).

Prather et al. (2009) suggest that because HVC_X neurons innervate striatal structures important for song learning and perception, the coding strategy employed by HVC_X neurons to represent vocal sequences, may have implications for learning and perception of speech in humans. “In the human brain, cortical neurons similar to HVC auditory–vocal neurons could transmit speech-related auditory and motor information to striatal regions implicated in speech development. Furthermore, auditory–vocal mirror neurons with properties similar to the HVC_X cells described here could bind sensory and motor features of distinct vocal gestures, providing an efficient substrate for rapid decoding and encoding of speech” (Prather et al., 2009).

Prather et al. (2009) point out that the division of continuously variable acoustic signals into discrete perceptual categories is a fundamental feature of vocal communication, including human speech. Despite this, the neural mechanisms involved have been poorly studied (Doupe and Kuhl, 1999). The authors point out that swamp sparrows learn their song notes by imitation, a feature of human speech, otherwise rare among animals. Behavioral

experiments had shown that male swamp sparrows use categorical perception to distinguish fundamental acoustic elements in their species-typical vocal repertoire. These note types (similar to phones in speech) are produced with considerable variation by different individuals but are grouped into natural categories.

As described above, the investigators had shown that the nucleus HVC contained a certain class of striatum-projecting neurons, HVC_X cells that respond to only one song type in the bird's repertoire, and song perception was shown to be impaired by lesions to the striatal portion of an AFP into which HVC_X cells project their axons. This sensorimotor correspondence was thus suggested by the investigators as being reminiscent of mirror neurons in the monkey cortex, hypothesized to be important in perception (Aronov et al., 2008).

Prather et al. (2009) recorded from antidromically identified HVC neurons in freely behaving male swamp sparrows, and presented each bird's song types through a speaker located near its perch as above. They presented each HVC cell a set of song stimuli comprising 5–11 variants of the primary song type, each differing only in the duration of a single replacement note in each trilled syllable, with duration of notes classified as category I to category VI. The procedure revealed that auditory responses of HVC_X neurons, but not interneurons were highly sensitive to changes in note duration. Robust responses were invariably evoked by stimuli containing replacement notes with durations that fell unambiguously into the same category as the target note in the natural song. In distinction, interneurons responded similarly when the replacement note was the same or a different category as the target note, indicating that "categorical responses to changes in note duration are shown by only one subset of auditory responsive HVC neurons, namely those that project to a striatal pathway that is important in song perception" (Prather et al., 2009).

Furthermore, the investigators were able to demonstrate geographically distinct populations of swamp sparrows obtained from northwestern Pennsylvania vs. upstate New York, in which perceptual categorical note boundaries differed from Category I to Category VI. This suggested that these differences may have been influenced by learning, and thus vary across populations. According to the investigators, the study provided the first evidence for neurons encoding perceptual information about a phonological feature of learned vocal behavior, specifically information about a categorical perceptual boundary. Variation in this perceptual boundary across swamp sparrow populations strongly suggested that both categorical perception and categorical neural responses in sparrows are affected by experience. This observation was linked by the authors to the role of neural experience in shaping human speech perception of categorical boundaries. Finally, the establishment that categorical responses are expressed by striatal projecting HVC_X neurons, but not by interneurons, was thought to closely parallel the activity of auditory afferents to HVC, where the highly selective auditory responses of HVC_X neurons required inhibitory sculpting through interneurons. This sculpting was thought to occur through a process of local inhibition, allowing context sensitivity over hundreds of milliseconds (Prather et al., 2009).

While birdsong may also be built on "mirror" foundations, the capacity of HVC neurons to "innervate striatal structures, important for encoding and decoding of learning and perception," as well

as the geographically distinct phonological features of northwestern vs. upstate New York swamp sparrows suggests an indexical function for these songs.

HUMAN LANGUAGE MATURATION

Werker and Tees (1999) point out newborn infants begin life with a remarkable sensitivity to the acoustic cues that signify different basic elements of speech. By measuring babies' sucking response to syllables such as /ba/ vs. /da/ or /ba/ vs. /da/, infants discriminated consonants most easily that actually occur in most of the world's languages. Werker and Tees point out that Japanese babies were able to hear the distinction between /r/ and /l/, but Japanese adults were unable to hear this distinction. According to Werker and Tees infants become relatively more sensitive to the phonetic characteristics of the native language, and also to the syllabic context in which that phonetic variation occurs (Werker and Lalonde, 1998; Werker and Tees, 1999).

The language-general perceptual sensitivities in newborns undergo a change and become more language-specific in the first year of life, thus preparing the infant for the ability to understand and speak his/her native language. During the first 14–15 months, infants learn to extract words from the speech stream, and to recognize word forms they have previously heard, and to associate words with objects. Coincident with the decline in non-native consonant (and vowel) discrimination seen by the end of the first year of life, the ability to co-ordinate two sources of information, such as phonetic detail, and position in a word is developed. The task for the next year of life is to construct a second-order system to effortlessly and efficiently use the medium of speech to map to meaning. With the establishment of a new level of representation, a discontinuity is produced (Werker and Lalonde, 1998). This discontinuity suggests a transformation from iconic mimicry to indexical representation, possibly similar to that which occurs in birdsong.

The decline in non-native speech perception at the end of the first year of life, accompanied by the improvement in native speech perception has been found to be predictive of later language development (Kuhl et al., 2008). According to Kuhl et al. language and a "critical period" have long been of interest to language scientists (Eimas et al., 1971; Werker and Lalonde, 1998; Kuhl et al., 2008). Conboy et al. (2008) suggested that native language perceptual abilities are associated with cognitive control abilities, which play a specific role in the ability to disregard irrelevant phonetic information, while maintaining attention to relevant information. Using a conditioned head-turn test of native and non-native speech sound discrimination and non-linguistic object retrieval tasks, sequencing attention, and inhibitory control, the investigators showed that native speech discrimination was positively linked to receptive vocabulary size, but not to cognitive control tasks, whereas non-native speech discrimination was negatively linked to cognitive control scores, but not to vocabulary size. The results suggested specific relationships between the development of native language, speech perception, and vocabulary (Conboy et al., 2008).

Kuhl et al. (2008) point out that studies of the maturation of the human auditory cortex show that between the middle of the first year of life and 3 years of age, there is a maturation of axons

entering the deeper cortical layers from the subcortical white matter; and neurofilament-expressing axons appear for the first time in the temporal lobe, with projections to the deep cortical layers of the brain, providing the first highly processed auditory input from the brain stem. The temporal coincidence between this cytoarchitectural change and infants' phonetic learning provides a possible maturational factor in the opening of a critical period for phonetic learning (Moore and Guan, 2001).

The concept of transition between developing brain and native language phonetic ability, as well as the associated concept of a developmental discontinuity to a second-order representational system indicates a possible basis for the understanding of the importance of language in development. The distinction made by Prather et al. (2008), between visual-motor and auditory-vocal "mirror" neurons may be important for human speech development. While the significance of cortico-thalamic-striatal (CTC) circuits in visual-motor development is well-described, a possibly analogous "Broca-Wernicke" circuit, in which auditory-vocal mirror neurons play a part is suggested. In both cases, an early process or sculpting of visual representational symbol in the PFC and phonetic category in language circuits, needs to occur prior to the stabilization of goal oriented capacity.

PFC DEVELOPMENT

According to Fuster, the PFC is phylogenetically one of the latest cortices to develop, having attained maximum relative growth in the human brain (Fuster, 2001). Fuster states that by myelogenic and synaptogenic criteria, the PFC is clearly late-maturing, and that the human prefrontal areas do not attain full maturity until adolescence. Fuster describes the lateral PFC as the neural substrate for the cognitive functions that support the temporal organization of behavior.

"To conduct its executive functions, the lateral PFC interacts with subcortical structures and with other parts of the association cortex. A cardinal function of the lateral PFC is the temporal integration of information for the attainment of prospective behavioral goals [...] there is evidence indicating that activation is maintained through recurrent circuits between PFC cells and posterior cortex [...]. It is served by two complementary and temporally symmetric functions: working memory and preparatory set. Both work together in every sphere of action, including speech" (Fuster, 2001).

Thus, Fuster suggests an analogous recurrent process is important for both working memory and for speech, which is important in understanding the developmental importance of speech and language development in early childhood. In the mature brain, working memory is thought to depend on an intact dorso-lateral PFC (DLPFC; Fuster, 2001). According to Tau and Peterson (2010), rudimentary working memory capacities have been observed in infants as young as 6 months of age, but performance on Piaget's A not B task (retrieval of a hidden object after a delay) is not in place till 9 months, and is not solidly in place for difficult tasks till middle childhood.

Miller and Cohen (2001) describe the PFC as having the properties required to achieve top-down behavioral control. These include the ability to maintain its activity robustly until a goal

is achieved, and second to have interconnections with all sensory systems, cortical and subcortical motor systems, and with limbic and midbrain structures involved in affect, memory, and reward. Thus the lateral and mid-dorsal PFC receives visual, somatosensory, and auditory information from the occipital, temporal, and parietal cortices. The dorso-lateral area 46 is connected with premotor areas that send connections to primary motor areas and the spinal cord, as well as cerebellum and superior colliculus. There are also dense interconnections with basal ganglia.

Miller and Cohen (2001) also point out that the PFC neurons are both individually selective and others bimodally selective for sensory cues, but in addition PFC neural activity is able to represent rules required to perform a particular task. The Miller and Cohen model requires feedback signals from the PFC to reach throughout the brain. Miller et al. (1996) were able to show that monkeys were able to maintain a working memory of a rewarded stimulus over time, and that target-specific activity appeared simultaneously in the PFC and parietal cortex. While other brain areas can sustain activity up to several seconds, the PFC is distinguished by the ability to sustain such activity in the face of intervening distractions (Hopfield, 1982).

Thus the PFC exhibits sustained activity that is robust to interference: multimodal convergence and integration of behaviorally relevant information; feedback pathways that can exert biasing influences on other structures throughout the brain; and ongoing plasticity that is adaptive to the demands of new tasks. This specialization is optimal for a role in the brain-wide control and coordination of processing. The mechanisms responsible for updating representations in the PFC must be responsive to changes in the environment, as well as resistant to updating irrelevant changes. Miller and Cohen hypothesized that dopamine (DA) might play an important role in this gating function. They suggested that dual concurrent influences on midbrain DA allow the system to learn while it gates, and where a DA-mediated gating signal leads to a successful behavior, its concurrent reinforcing effects will strengthen the association of the signal with cues representing the pattern of activity that produced the behavior. Thus this self-organizing boot-strapping mechanism averted the invocation of a "homunculus" to control behavioral selection (Miller and Cohen, 2001).

Goldman-Rakic et al. (1990) have described the anatomical overlap of different mono-aminergic receptors in the same cortical strata, suggesting that there may be families of receptors linked by localization on common targets. Arnsten points out that although Goldman-Rakic (1994) used spatial working memory as a model system for examining functional circuitry, she proposed that these principles applied to other sensory and affective domains, and described the process as "representational knowledge within parallel processing streams." According to Arnsten, Goldman-Rakic spoke of PFC network activity as a fundamental contribution to mind, and the disruption of this process as a primary contribution to thought disorder in mental illness. "She used the term working memory to describe a building block of cognition: the ability to represent information no longer in the environment through recurrent excitation of pyramidal cells with shared stimulus properties" (Arnsten, 2007). Arnsten described the role of the PFC in working memory, as applying representational knowledge

to inhibit inappropriate action, thought, and feelings, as well as inhibiting responses to distracting stimuli. However, Arnsten and Goldman-Rakic (1985) were able to show that “many effects formerly attributed solely to DA, involved both NE and DA actions.” According to Arnsten (2007), both DA and NE exhibit an inverted-“U” dose/response, where either too little or too much arousal impairs working memory.

An important distinction outlined by Arnsten relates to the location of D1 and alpha-2A signaling mechanisms on dendritic spines. She points out that under optimal neurochemical conditions, moderate levels of NE engage alpha-2A receptors, and increase signals, whereas moderate levels of DAD1 receptor stimulation decrease “noise.” These beneficial effects of alpha-2A vs. DAD1 arise from opposing effects on cAMP signaling, where alpha-2A stimulation inhibits, while DAD1 activates cAMP production. Thus D1/alpha-2A signaling appears to have an important *representational* role in visuo-spatial working memory (Arnsten and Goldman-Rakic, 1985). The demonstration by the work of Goldman-Rakic and Arnsten (Arnsten and Goldman-Rakic, 1985; Goldman-Rakic, 1994; Arnsten, 2007) on the importance of symbolic representational capacity in the human prefrontal cortex for working memory has implications for language development. It is likely that human language is distinguished from primate indexical reference by its capacity to recursively encode, incorporate, and combine visual and auditory symbols in working memory, basic for human language functions.

DISCUSSION

The classification of referential capacity in terms of hierarchical iconic, indexical, and symbolic levels allows an understanding of the hierarchical nature of mirror, birdsong, and human language capacity. Birdsong studies have revealed that the HVC was not necessary for the generation of subsong or early babbling, whereas the generation of adult “syllabic” complex stereotyped motifs required an inhibitory output from the HVC to brainstem motor nuclei, but nonetheless remain stereotypic and indexical. From the birdsong analogy, activity during the subsong or “babbling” stage represents a sculpting of categorical native phoneme recognition, while mature song requires stabilization of HVC_X–RA song circuits. Similarly, sculpting of human “mirror” neurons may play a part in the development of capacity for symbolic representation in the DLPFC.

REFERENCES

- Arbib, M. A. (2010). “The mirror system hypothesis,” in *Action to Language via the Mirror Neuron System*, ed. M. A. Arbib (Cambridge: Cambridge University Press), 3–47.
- Arbib, M. A., and Bota, M. (2010). “Neural homologies and neurolinguistics,” in *Action to Language via the Mirror Neuron System*, ed. M. A. Arbib (Cambridge: Cambridge University Press), 136–167.
- Arbib, M. A., and Rizzolatti, G. (1997). Neural expectations: a possible evolutionary path from manual skills to language. *Commun. Cogn.* 29, 393–423.
- Arnsten, A. F. T. (2007). Catecholamine and second messenger influences on prefrontal cortical networks of “representational knowledge”: a rational bridge between genetics and the symptoms of mental illness. *Cereb. Cortex* 17(Suppl. 1), i6–i15.
- Arnsten, A. F. T., and Goldman-Rakic, P. S. (1985). Alpha-2 adrenergic mechanisms in prefrontal cortex associated with cognitive decline in aged nonhuman primates. *Science* 230, 1273–1276.
- Aronov, D., Andalman A. S., and Fee, M. S. (2008). A specialized forebrain circuit for vocal babbling in the juvenile songbird. *Science* 320, 630–634.
- Baddeley, A. D. (2003). Working memory: looking back and looking forward. *Nat. Rev. Neurosci.* 4, 829–839.
- Conboy, B., Sommerville, J., and Kuhl, B. (2008). Cognitive control factors in infant speech perception at 11 months. *Dev. Psychol.* 44, 1505–1512.
- Deacon, T. W. (1997). *The Symbolic Species: The Co-Evolution of Language and the Brain*. New York: WW Norton, 69–90.
- Doupe, A. J., and Kuhl, P. K. (1999). Birdsong and human speech: common themes and mechanisms. *Annu. Rev. Neurosci.* 22, 567–631.
- Eimas, P. D., Einar, R., Jusczyk, P., and Vigorito, A. (1971). Speech perception in infants. *Science* 171, 303–306.
- Fitch, W. T. (2005). Protomusic and protolanguage as alternatives to proto-sign. Commentary: from monkey-like action recognition to human language: an evolutionary framework for neurolinguistics. *Behav. Brain Sci.* 28, 132–133.
- Fuster, J. M. (2001). The prefrontal cortex—an update: time is of the essence. *Neuron* 30, 319–333.

The present review suggests that similar processes occur in human language development, where categorical native phoneme recognition is sculpted in a Wernicke/Broca circuit en route to sequential language capacity. Studies of human language development reveal an analogous early babbling stage, during which there is a transition from generalized to native phonemic usage, with a subsequent childhood transition to second-order representational capacity. The capacity for association visual and auditory symbols in working memory providing a basis for human language development. Adequate PFC functioning appears critical for not only mature reasoning, but also involves behavioral functions, including inhibition of task-irrelevant behaviors, processing of affect, motivation, and reward attainment by virtue of connections with wide-ranging cortical centers. A consequence of such deficits in PFC development is an incapacity for sequential reasoning, lack of affect regulation, a lack of capacity for working on sustained goal achievement, and a tendency for impulsive and repetitive behaviors, under either environmental, or subcortical control. It can thus be argued that the process of development is closely dependant on adequate PFC development, and many if not most behavioral syndromes of childhood reflect deficits in cortical development. Importantly this includes language development, where auditory-vocal “mirror” neurons may have an important role in the transition from babbling to goal-directed language (Levy et al., 1987; Levy and Hobbes, 1989).

SUMMARY

Three fundamental forms of cognitive reference: iconic, indexical, and symbolic are described in relation to mirror systems, birdsong, and human language. The process of human development is closely dependant on adequate PFC development, (and many if not most behavioral syndromes of childhood reflect deficits in cortical development). Importantly this includes language development, where auditory-vocal “mirror” neurons may have an important role in the transition from babbling to goal-directed language. While the significance of CTC circuits in visual-motor development is well-described, a possibly analogous “Broca–Wernicke” circuit, in which auditory-vocal mirror neurons play an initial part is suggested. In both cases, an early process or sculpting of visual and auditory representational symbols in the PFC and language circuits occurs as a basis for human language.

- Goldman-Rakic, P. (1994). "The issue of memory in the study of prefrontal functions," in *Motor and Cognitive Functions of the Prefrontal Cortex*, ed. A. Thierry (New York: Springer-Verlag), 112–122.
- Goldman-Rakic, P. S., Lidow, M. S., and Gallager, D. W. (1990). Overlap of dopaminergic, adrenergic and serotonergic receptors and complementarity of their subtypes in primate prefrontal cortex. *J. Neurosci.* 10, 2125–2138.
- Hopfield, J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci. U.S.A.* 79, 2554–2558.
- Iacoboni, M., and Dapretto, M. (2006). The mirror neuron system and the consequences of its dysfunction. *Nat. Rev. Neurosci.* 2, 942–951.
- Kuhl, P., Conboy, B., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., and Nelson, T. (2008). Phonetic learning as a pathway to language: new data and native language magnet theory expanded (nlm-e). *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363, 979–1000.
- Levy, F., Horn, K., and Dalglisch, R. (1987). Relation of attention deficit and conduct disorder to vigilance and reading lag. *Aust. N. Z. J. Psychiatry* 21, 242–245.
- Levy, F., and Hobbes, G. (1989). Reading, spelling and vigilance in attention deficit and conduct disorder. *J. Abnorm. Child. Psychol.* 17, 291–298.
- Miller, E. K., and Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* 24, 167–202.
- Miller, E. K., Erickson, C. A., and Desimone, R. (1996). Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *J. Neurosci.* 16, 5154–5167.
- Moore, J. K., and Guan, Y. L. (2001). Cytoarchitectural and axonal maturation in human auditory cortex. *J. Assoc. Res. Otolaryngol.* 2, 297–311.
- Peirce, C. S. (1931–1958). *Collected Papers*, Vol. 1–6, eds Hartshorne and Weiss (Cambridge, MA: Harvard University Press).
- Prather, J. F., Peters, S., Nowicki, S., Anderson, R. C., Peters, S., and Mooney, R. (2009). Neural correlates of categorical perception in learned vocal communication. *Nat. Neurosci.* 12, 221–228.
- Prather, J. F., Peters, S., Nowicki, S., and Mooney, R. (2008). Precise auditory-vocal mirroring in neurons for learned vocal communication. *Nature* 305–312.
- Rizzolatti, G., and Arbib, M. A. (1998). Language within our grasp. *Trends Neurosci.* 21, 188–194.
- Rizzolatti, G., and Craighero, I. (2004). The mirror-neuron system. *Annu. Rev. Neurosci.* 27, 169–192.
- Tau, G., and Peterson, B. (2010). Normal development of brain circuits. *Neuropsychopharmacology* 35, 147–168.
- Werker, J. F., and Lalonde, C. E. (1998). Cross-language speech perception: initial capabilities and developmental change. *Dev. Psychol.* 52, 672–683.
- Werker, J. F., and Tees, R. C. (1999). Influences on infant speech processing: toward a new synthesis. *Annu. Rev. Psychol.* 50, 509–535.

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 10 May 2011; accepted: 21 December 2011; published online: 11 January 2012.

Citation: Levy F (2012) Mirror neurons, birdsong, and human language: a hypothesis. *Front. Psychiatry* 2:78. doi: 10.3389/fpsy.2011.00078

This article was submitted to *Frontiers in Child and Neurodevelopmental Psychiatry*, a specialty of *Frontiers in Psychiatry*. Copyright © 2012 Levy. This is an open-access article distributed under the terms of the Creative Commons Attribution Non Commercial License, which permits non-commercial use, distribution, and reproduction in other forums, provided the original authors and source are credited.