

1. INTRODUCTION

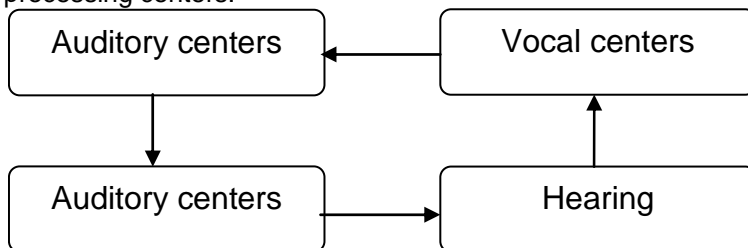
1.1. Vocal Learning

Vocal learning – the ability to imitate complex sounds – is a rare form of learning that is known to occur in humans and 3 groups of birds (parrots, hummingbirds, songbirds); as well, there is some evidence for vocal learning in dolphins, whales, and bats (reviewed in: Janik and Slater, 2000; Jarvis, 2004; Snowden and Hausberger, 1997). Vocal learning is a behavior that requires a non-trivial conversion from sounds heard by an individual to sounds produced by that individual. The time gap between hearing sounds and producing similar sounds might be remarkably long. For example, a juvenile song sparrow listens to song of adult males in the spring, but he imitates those songs only during the next spring, even if kept isolated after hearing the songs of his tutor (Marler, 1964). Therefore, vocal learning (in birds) involves the formation of unique, and long lasting memories, often referred to as an 'auditory template' for imitation (Konishi, 1965). The auditory template is presumably mapped into a set of motor commands that produce an acoustic signal similar to the sensory template – how exactly sensory input is converted to appropriate acoustic output is not known. Further, despite 50 years of efforts, “auditory template memories” are still elusive and theoretical: they have never been conclusively localized in any species (see however Phan, Pytte, Vicario, 2006; Troyer and Doupe, 2000). However, “template memories” for vocal learning are not dormant before vocal production starts to emerge: studies in human infants show that a great deal of perceptual tuning to the linguistic environment occurs in the first year of life prior to the emergence of spoken language (Kuhl, Williams, Lacerda, Stevens, Lindblom, 1992; Polka and Werker, 1992).

What sort of diachronic and synchronic interactions might occur between the motor and perceptual systems during vocal learning? Konishi suggested half a century ago (1965), for vocal imitation to occur there must be some feedback between motor experimentation and sensory perception. For example:

Figure 1-1

Vocal learning requires interaction between vocal centers, motor output, sensory input, and auditory processing centers.



This interaction between motor and sensory systems is especially interesting in the developing “vocal learner”. How do developmental processes influence, enhance, or restrict, vocal learning?

1.2. Developmental vocal learning

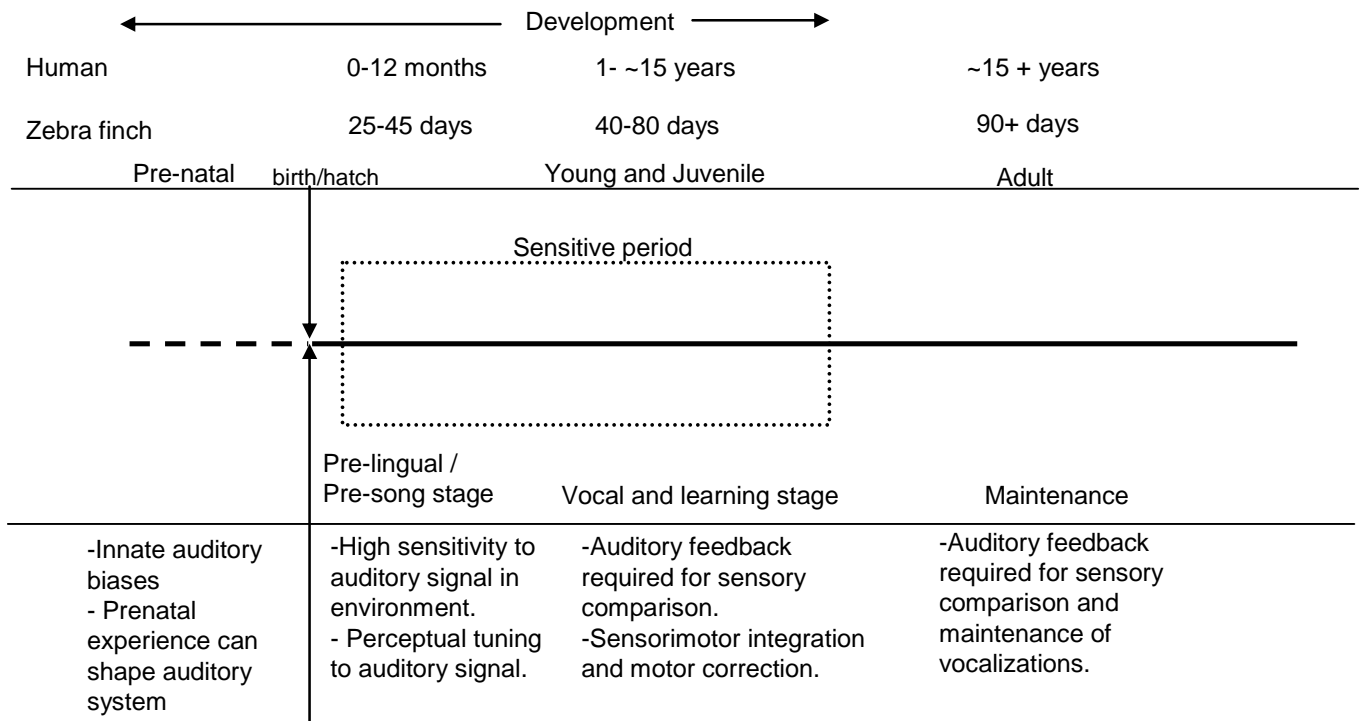
With respect to ontogeny, there are three types of vocal learning: i) developmental vocal learning (close-ended, or age-limited) occurs only or mostly during development (as in humans), ii) seasonal vocal learning usually occurs in association with reproduction (e.g., twice a year in the Canary), and iii) open-ended vocal learning occurs throughout life (e.g., in parrots, and to some extent in humans) (Janik and Slater, 2000). Developmental vocal learning is most prevalent among these. It occurs in both humans learning language and some of the songbirds (*Passerimorphes*, suborder Oscines). This dissertation focuses on developmental vocal learning in humans and songbirds, with particular attention to the songbird zebra finch (*Taeniopygia guttata*), a species commonly used to study mechanisms of vocal learning.

Developmental vocal learning requires:

- Auditory signal: sensory input to the infant or juvenile bird, usually species-specific vocalizations;
- Sensitive period: developmental time-window during which the infant or juvenile bird is sensitive to the auditory signal;
- Auditory feedback: ability for the infant or juvenile bird to hear its own vocalizations and compare these to internal and external auditory signals.

Figure 1-2

Schematic of vocal learning process in humans and zebra finches. Fundamental requirements of auditory signal, sensitive period and auditory feedback are similar although specific contributions of each component and the time-scale of learning is different between the species.



A major difficulty in understanding the developmental interactions between sensory and motor processes involved in vocal learning is the lack of tools for performing comparative studies that would allow us to take advantage of different model systems. For example, in humans sensory changes and perception can be studied in early stages of development, but the sensory environment cannot be controlled. In songbirds, the developmental environment can be controlled, but the brain imaging techniques used in humans have no parallel in songbirds. This dissertation focused first on developing such imaging techniques for songbirds, and then on using them in a way that might complement studies in humans – namely controlling the environment during vocal learning – to ask, at the most fundamental level, how developmental experience: social, auditory and vocal, might affect sensory processes. We focused on the sensory aspects of developmental vocal learning in the songbird zebra finch (*Taeniopygia guttata*), a species commonly used to study mechanisms of vocal learning.

1.2.1. Auditory Signal and innate biases:

The natural auditory arena is always filled with a variety of sounds. How can the animal (or infant) decide which sounds should be imitated? It is often assumed (but rarely tested) that most vocal learners have a mechanism allowing them to filter out many environmental sounds and focus on species-specific vocalizations based on acoustic features.

In humans

In human newborns, the extent of tuning to the speech signal is remarkable. Newborns appear to be born with strong preference for listening to speech and detecting speech sounds and language (Vouloumanos & Werker, 2007). Part of these abilities may be due to the intact fetal auditory system that detects internal (maternal visceral) and external sound from approximately 20 weeks gestational age (with attenuation up to 20 dB) (Johansson, 1964). Prenatal auditory experience with species-specific and language specific vocalizations accounts for some of auditory preferences observed in newborns: preference for mother's voice (DeCasper & Fifer, 1980), prosodic pattern of maternal language (Moon, Cooper, Fifer, 1993), and even familiar words and passages (DeCasper & Spence, 1986). [Not all aspects of hearing are mature in the newborn: absolute auditory thresholds are elevated (Northern & Downs, 1984)]; and frequency selectivity, detection of sound-in-noise, sound localization, and auditory attention do not mature until months or years after birth).

The earliest – prenatal – auditory experience does not account for the universal speech perception of infants, and their ability to discriminate fine-grained phonetic contrasts which are not audible in utero, and also not in the immediate language environment at birth. Eimas' (1971) showed that in the first 4 months of life, before the infant is able to produce learned vocalizations, the sensory system is already tuned to spoken language and the infant can discriminate speech contrasts, for example, minor acoustic (and phonemic) cues in syllable pairs such as timing differences in voicing of the vowel component of the syllables [pa]:[ba] (Eimas, Siqueland, Jusczyk, Vigorito, 1971;). Moreover, this discrimination ability appears to be adult-like, categorical and universal for languages: when tested with pairs of consonants with equal-step acoustic changes along a continuum, infants show similar perceptual

boundaries to adult speakers of various languages (Aslin, Pisoni, Hennessey, Perey, 1981; Werker and Tees, 1981;). Newborns are equipped with categorical perception, but is it really an innate bias specialized for speech? Additional studies showed that infants have categorical perception of non-speech sounds (Jusczyk, Pisoni, Walley, and Murray, 1980) and that non-human mammals exhibit categorical perception of speech sounds (Kuhl and Miller, 1975). These findings suggest that the initial state of “universal speech perception” may be based on general features of the mammalian auditory system (review Doupe and Kuhl, 1999). These general auditory biases and prenatal auditory experiences are further strongly and rapidly molded by the auditory experience of the infant: the infant becomes tuned to the speech sounds in its language environment (language spoken by parents and care-givers). By one year infants have lost much of their “universal” perception but fine-tuned their discrimination of speech sounds and linguistic features of what will be their native spoken language (Gilbert, Humphrey, Tees, 1981; Werker and Tees, 2002).

The studies above, and many more, have used (now) well-established behavioral techniques to observe infant auditory preferences and perception. These methods are non-invasive and rely on natural infant behavior, such as, sucking on a pacifier, and interest in novelty (review, Borden, Harris, Raphael, 2003). For example, in response to a novel sound the infant will increase rate of sucking, in this way discrimination of acoustically similar sounds can be tested. Only in the last 10-15 years has it been possible to study the infant brain more closely, and the changes in brain activity that are the basis for the “universal” speech perception and the perceptual tuning that lead to their remarkable ability to acquire speech.

In songbirds

What is known about initial auditory *perceptual* state and effect of pre-hatch auditory experience in songbirds? The zebra finch also hatches in an altricial state, in which perceptual testing involving behavioral tasks are difficult to carry out (although investigation of begging responses should be possible even in very young chicks). However, tests of auditory function (Auditory Brainstem Responses) in the very young zebra finch indicate that the 10 day post-hatch bird has immature hearing thresholds, 40 dB above adult thresholds (at approximately 70-100 dB SPL) (Amin, Doupe, Theunissen, 2007). Chicks of this age cannot hear much of the adult zebra finch vocalizations which are typically of low intensity. Hearing sensitivity matures by 20 days post-hatch – shortly after fledging – and at this age ABR thresholds are already adult-like (Amin, et al., 2007). In the earliest stage of zebra finch development (pre- 20 days post-hatch), it is likely that auditory perception is stable, as very little signal penetrates the auditory system.

In absence of behavioral measures of early auditory perception, initial proposals of innate perceptual tuning in songbirds (the “innate auditory template”) were based on observations of the outcome of early sensory exposure on production, rather than by direct estimates of perception. Birds deafened early in life later produced song that contained species-specific features indicating an innate motor program for song (swamp sparrows; Marler and Sherman, 1983; zebra finches: Price 1979;). Hearing birds learned conspecific rather than heterospecific song when exposed to both during

development, suggesting that conspecific song is a more salient auditory signal for these birds (sparrows) (Marler and Sherman, 1983). Early studies (Konishi, 1965; review Marler and Slabbekoorn, 2004) clearly show that there is an innate component (motor and sensorimotor) to song learning and vocal production in a number of species (contribution of innate component varies widely by species). More recent studies have used operant conditioning using food reward (e.g. animal hops to perch, or pecks on a key to choose a reinforced song) to study auditory perception in adults (Dooling, 1973; Dooling and Searcy, 1980). Results here suggest a sensory or perceptual bias in both males and females for species-specific vocalizations, but in birds who have lots of auditory experience with species-specific vocalizations (similar to the infants at 1 year) (Cynx and Nottebohm, 1992; Riebel, 2000). In a few cases song preference has been tested in juveniles zebra finches naïve to song (i.e. raised by females until day 28-53) and physically able to perform perceptual tasks; these birds too show a general preference to listen to conspecific over heterospecific song, (Braaten and Reynolds, 1999; Braaten, Petzoldt, Colbath, 2006;). It appears that in many songbirds there is something about species-specific vocalization that is innate or preferred on both the sensory and the motor side, but the details of the innate auditory filter or template are still missing: where it is located in the brain, and its specific acoustic features (although see Woolley and Rubel, 1999).

In songbirds, early auditory environment can be controlled and there is conclusive evidence that some of the sensory filters are indeed innate. These “innate auditory templates” in the songbird (often called “auditory biases” in human infants) make the juvenile bird or infant especially receptive to the “appropriate” auditory signal: species-specific vocalizations in their environment (Dooling and Searcy, 1980; Vouloumanos and Werker, 2004). The particular sounds that the animal or infant eventually selects as a template for imitation are often the vocalizations of the individuals caring for them, e.g., the parent. However, we do not know if and how the early exposure to songs might shape perception in songbirds, in a way that might further “help them decide” which sounds should be imitated. In both infants and songbirds the sensory system may be specialized to process specific environmental sounds or ‘tuned’ to these sounds, and in addition, the perceptual system is further modified by exposure to sounds. The interesting questions are of the mechanisms, the time-frame and the nature of the experience that shapes the perceptual system.

1.2.2. *Sensitive period:*

A sensitive period in the development of any organism is a window of developmental-time in which the effect of environmental influences (sensory input) is stronger than at other time in life (Aslin, 1981). During a sensitive period, sensory experience may directly affect the wiring of the developing brain, by modifying neural connections and cytoarchitecture, which consequently influence sensory systems, perception, and motor function (Bottjer, 2002; Hubel & Wiesel, 1978).

The sensitive period in developmental vocal learning is the time during which the learner is most receptive to the auditory signal, and the sensory system becomes tuned to the environment specific auditory signal (Janik and Slater, 2000). During the sensory phase of the sensitive period auditory template memories are acquired; these template memories are then converted into vocal sounds in the

sensory-motor phase of the sensitive period (Konishi, 2004). These phases often overlap, or appear to be extended, and the existence of a rigid sensitive phase has been contentious in the history of both human and songbird research (Doupe and Kuhl, 1999). For example, white-crowned sparrows imitate sounds heard from a playback only during the sensitive period, but given social interaction with a live singing tutor adult birds can acquire a song template even after the sensory period (Baptista & Petrinovich, 1986, but see Nelson, 1998). In humans the sensory phase for spoken language potentially lasts years (for acquisition of subtle aspects of speech and language; i.e. complex syntax, vocabulary and phonotactics) (Berko-Gleason, 1993). But the most dramatic perceptual tuning takes place within a very short period of time early in life – the first 12 months – before a spoken utterance has been acquired (Kuhl et al., 1992; Polka and Werker, 1994). Individual variation in perception during this early period is associated with the development of language skills later in life (Tsao, Liu, Kuhl; 2004), but deaf infants who have no exposure to speech from 0-12 months of age are still capable of developing spoken language later in life, e.g., once a cochlear implant is in place. In such cases, the fine quality of speech is correlated with the age of implantation, but the basic ability to acquire spoken language is not entirely lost after the sensory phase of the first year of life (Yoshinago-Itano, Sedey, Coulter, Mehl. 1997).

In both humans and close-ended songbirds the end of the sensitive period is associated with sexual maturation. In humans this generally refers to the observation that children learn first and second languages relatively quickly, without implicit instruction, and completely, without weakness in grammar or phonology but the ease of speech and language learning declines with the onset of adolescence (Birdsong, 1999). Adults learning a second language, with or without implicit instruction, generally, have more difficulty learning the language and attaining native-like proficiency in grammar and phonology (Johnson & Newport, 1989). There are exceptions to this phenomenon, and there is much dispute about the exact nature of the, sensitive, or 'critical period' for speech and language in humans (Birdsong, 1992). In addition, second language learning in adults appears to be characteristically different from first language learning in children, and factors such as goals of learning the language, linguistic awareness, cultural identity may influence the ultimate language outcome of adult language learners (but are not as influential in children) (Johnson & Newport, 1989). For obvious ethical reasons there is limited information on *first* language learning after puberty and apparent closure of the sensitive period. The few individuals with normal hearing but deprived of language input during childhood are confounded by psychological trauma (i.e. Genie, 1970's) (Curtiss, 1974). In these few case studies language (i.e. first language) was *not* acquired after puberty, suggesting that language input during the sensitive period is required for complete language acquisition later in life (Curtiss, 1974).

In the zebra finch the closure of the sensorimotor phase of the sensitive period is generally marked by song crystallization; stabilization of note structure and sequence, and increased stereotypy of song (Tchernichovski, 2001). Song crystallization can be manipulated, e.g. delayed with early castration, or delayed with lack of sensory input, but not indefinitely (Korsia and Bottjer, 1991). Development, or

maturation, of the bird eventually results in song crystallization regardless of the song quality: song crystallized in isolate birds, and tutored birds with poor imitation (review: Mooney, 1999).

In both human and songbirds, sexual maturity marks the end of the sensitive period for developmental vocal learning (Doupe & Kuhl, 1999); at this point the juvenile is producing adult-like vocalization of the song or language he has learned and the vocalizations are stable and stereotyped.

1.2.3. Auditory feedback

The ability for an animal or human learner to hear its current vocalizations is essential for developmental vocal learning, as well as to the maintenance of learned vocalization. Real-time influence of auditory feedback has been demonstrated in both humans (Doupe and Kuhl, 1999) and songbirds (Leonardo & Konishi, 1999). Humans use real-time auditory feedback to tune and guide their spoken language; delay of auditory feedback within a 200ms time-window induces dysfluency in normally fluent individuals (Lee, 1950). Permanent inability to monitor one's own speech and language production results in nonsense, rambling and jargon speech (Caspari, 2005). It is difficult to test prolonged disruption of auditory feedback during vocal learning in infants and children (i.e. during the sensitive period), but short-term delayed auditory feedback has an immediate effect on intensity of cry or vocalization in infants and toddlers (though these changes are not always in the same direction) (reviewed in Reese & Lipsitt, 1984). Lack of auditory feedback may be the cause of reduced acoustic variability and frequency in the vocal babbling of deaf infants, but in this population it is difficult to distinguish the effect of no auditory feedback versus lack of auditory input (Oller & Eilers, 1988). Adult examples of disrupted auditory feedback (lesion of the auditory language association area) result in an immediate and significant effect on vocal production (Villacorta et al., 2007).

After a songbird has acquired a song template it can perform the sensory-motor conversion and imitate the template song in complete isolation. Strikingly, the bird must hear himself during that isolation period in order to develop an accurate imitation of the song model he has heard (review Brainard and Doupe 2000; Konishi, 1965). This finding has led to numerous studies examining how the ability for the learner to hear its current vocalizations allows imitation. One hypothesis is that the sensory information is compared to the sensory memories or sensory template of the auditory signal and the resulting "error" drives changes in motor output (Konishi, 1965). In either early or late exposure to a song model, auditory feedback is required for the bird to produce prototypical syllables of the song template, or to change existing vocalizations (in mid-development) to match the song model. Song develops in the direction of the auditory template incrementally, as the bird compares self-produced vocalizations to the auditory template, and adjusts his motor output accordingly (Konishi, 1965).

In contrast to the early view of auditory feedback as required only for vocal learning, it is now clear that auditory feedback is equally important in adults (Leonardo & Konishi, 1999). In both adult humans and songbirds, auditory feedback is essential for maintenance of stable vocalizations and perturbations in auditory feedback after vocal learning produce varying degrees of real-time change in stable vocalizations (Sakata and Brainard, 2006; Villacorta et al., 2007).

1.3. Auditory and vocal systems functional anatomy in a nutshell

In this dissertation we look at brain activation in birds, how can this be relevant to humans? Brain structures in aves and mammals are not homologous and the brains of humans and zebra finches in particular are quite disparate: 1) there is a great size difference (zebra finches: ~2cm in diameter); 2) human telencephalon (forebrain) contains a highly developed neocortex: the highly convoluted, laminar (six-layered) cerebral cortex (gyri-sulci increase surface area 10-fold, further accentuating the size difference between zebra finch and human brains), whereas it has only recently been recognized that a significant, rather than minor, portion of the avian brain develops from the telencephalon; 3) avian brain is *not* laminar, but nucleated, and neuronal arrangement and interactions are quite different from those in mammals (Butler, Manger, Lindahl, Arhen, 2005; Jarvis, et al. 2005). Despite these gross differences in brain structures there are newly recognized similarities in functional parts of the brain and well-known similarities in specific behaviors between humans and zebra finches, i.e. developmental vocal learning (Güntürkün, 2005). A shift in to thinking about function over anatomy, and new techniques that allow greater detail in observing cytoarchitecture and neural connections have led to a new view of the avian brain, and revision of the avian brain nomenclature (Farries, 2004; Reiner, Perkel, Mello, Jarvis, 2004). In the 1970's nearly three quarters of the avian brain had been classified as 'striatal' (subcortical – primarily because this area was nucleated rather than laminar, so anatomically resembled mammalian subcortical rather than cortical structures). It is now accepted that a large part of the avian brain is functionally forebrain (though nucleated), receives abundant auditory and visual input from true subcortical regions (thalamus & basal ganglia), and performs sensory & motor functions reminiscent of the mammalian neocortex (Reiner, et al., 2004). The revised nomenclature suggests a relatively large neopallium (cerebrum) in the avian brain, strikingly similar to the large neocortex of the mammalian brain (Jarvis, et al., 2005).

The few groups of birds and mammals that have vocal learning - 3 groups of birds (songbirds, parrots, and hummingbirds) and 4 groups of mammals (cetaceans, bats, elephants and humans) - are phylogenetically distant, and vocal learning appears to have evolved independently in each (Jarvis et al., 2005). Despite evolutionarily different paths the fundamental behavior of vocal learning is maintained across these groups: all exhibit ability to modify vocal output based on auditory input. Obviously specifics of what is learned and how vary greatly: human spoken language is symbolic, creative and recursive; zebra finch song is simple and stereotyped; parrots exhibit pair duets (sequential calling), call convergence, and mimicry of exotic sounds) (Doupe and Kuhl, 1999; Jarvis, 2007). What common brain mechanisms could exist that support this vocal learning? Beyond brain structures – which are not homologous - are there common mechanisms for vocal learning across these groups? Several hints to such mechanisms have been revealed: at the molecular level the FOXP2 gene has been implicated in human language disorders and abnormal song development in zebra finches (White, Fisher, Geshwind, Scharff, Holy, 2006); at behavioral levels the tendency for vocal experimentation during the sensorimotor

phase of vocal development in both humans and zebra finches (Doupe & Kuhl, 1999; Oller, 2000; Tchernichovski, 2001).

zebra finch

Zebra finch is a recently domesticated species that has a short developmental period (sexual maturity at 80-90 days post-hatch). Their continuous, non-seasonal reproduction, rapid song development, and ability to tolerate invasive electrophysiology while singing, make them an excellent animal model for studying vocal learning. Sexual dimorphism exists in the zebra finch brain: males have song control nuclei, in females these nuclei are undeveloped or not present (Nottebohm and Arnold, 1976). The presence of song motor nuclei and connections between them are directly related to singing behavior: males learn song; females do not learn to sing, but select mates based on song complexity and singing performance (Zann, 1996). The description below is based on male zebra anatomy, Figure 1-3.

Sensory – In the songbird, much like in mammals, auditory signal is transferred from the periphery to the neopallium through multiple nuclei: second, third, and fourth order nuclei in the brainstem (Reiner et al., 2004; Theunissen, Woolley, Hsu, Fremouw, 2004) to auditory nuclei in the thalamus. The thalamic nuclei Ovoidalis (Ov) provides the major auditory input to the telencephalon, projecting to Field L, a complex structure and a region analogous to the mammalian auditory cortex. Field L has multiple functionally and structurally distinct subregions – L, L1, L2a, L2b, L3 – and projections to secondary auditory regions caudal nidopallium (NCM), caudal mesopallium (CM) and lateral caudal mesopallium (CLM) (Vates, Broome, Mello, Nottebohm, 1996). Functionally, Field L neurons are selective to natural sounds over synthetic sounds (Grace and Theunissen; 2003) but do not show preference for subjectively relevant stimuli: Bird's Own Song (BOS), Tutor song (TUT). Neurons in adjacent nuclei (NCM, CM) show increased selectivity to complex stimuli, such as BOS and TUT, suggesting that these areas (NCM, CM) perform higher level processing of the sound (Phan, et al., 2006). Nif (Nucleus Interfacialis) is a sensorimotor nucleus adjacent to the auditory processing areas, receiving thalamic input and auditory input (from CM) and projecting to the vocal nucleus, HVC (Theunissen, et al., 2004). Nif provides primary auditory input to the premotor nucleus HVC, whereas CM and Field L projections enter the HVC “shelf”, a region on the ventral border of HVC (review Theunissen et al., 2004). The input of auditory signals from HVC shelf to HVC proper, is not clear (Wang, Sakabe, Sakaguchi, 2001). There are multiple synapses and increasingly complex processing of auditory stimuli between the primary “auditory cortex”, Field L and the primary premotor nucleus of the song system, HVC. The sensorimotor nucleus Nif (selective auditory center) appears to be the interface between the auditory and motor systems (Amin, Grace, Theunissen, 2004)

Motor– Song production is mediated by the “motor pathway”, a circuit that extends from the thalamus, through nuclei in the neopallium, to motor neurons that are active during singing and respiration (Mooney, 2000). HVC is the primary nucleus in the song system, receiving ascending inputs from auditory nuclei and providing premotor input to robust nucleus of the arcopallium (RA) and the

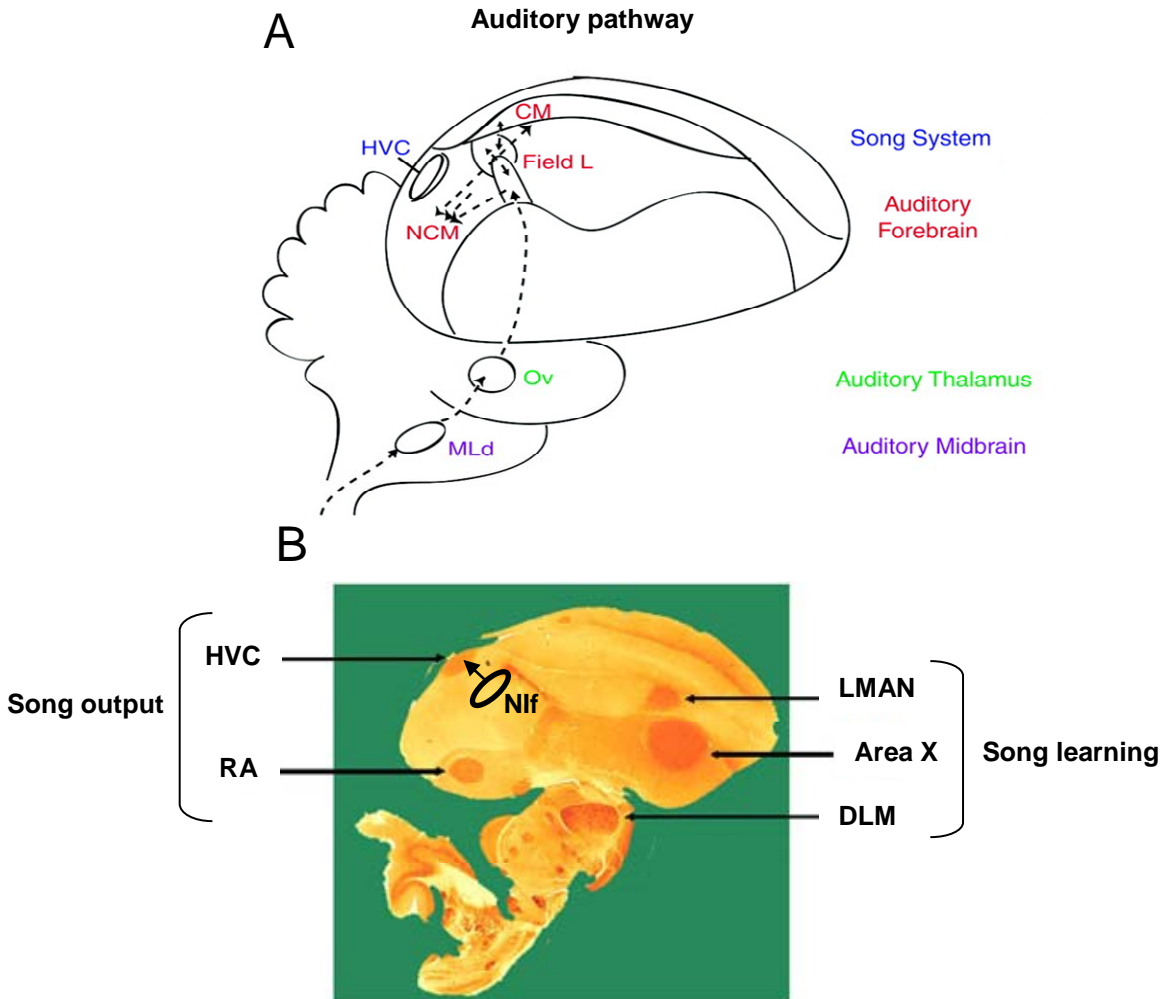
anterior forebrain, the site of the learning pathway (Mooney, 2000; Nottebohm, 1991). RA projects to vocal, respiratory, and laryngeal nuclei in the midbrain; tracheosyringeal motor neurons from the laryngeal nucleus innervate muscles of the vocal organ, the syrinx (Goller and Cooper, 2004). Neural activity in HVC and RA is synchronized with sound production by the syrinx (Hahnloser, Kozhenvnikov, Fee; 2002).

Learning– An additional pathway is specific to the behavior of song learning but does not appear to be necessary for song maintenance in adulthood (Solis and Doupe, 2000). The anterior forebrain nuclei form an indirect path from HVC to RA: Area X receives direct input from HVC and projects to the medial dorsal lateral nucleus in the thalamus (DLM). DLM projects to the lateral magnocellular nucleus of the anterior neopallium (LMAN) also in the forebrain, which feeds back to Area X and forward to RA (review, Bottjer, 2002; Brenowitz, Margoliash, Nordeen, 1997).

Figure 1-3

Sensory, Motor and Learning pathways in the zebra finch song system.

A. Different levels of the auditory pathway: Field L receives thalamic auditory input, auditory signals from Field L are sent to L subregions and to the secondary auditory areas (NCM and CM complex). Nlf (not shown) receives input from secondary auditory areas and provides the primary auditory input to the song system (HVC), whereas NCM and CM project to the shelf surrounding HVC (presumed indirect input to HVC). Modified with permission from Amin, et al., 2007. **B.** Song system motor nuclei: HVC the premotor song nucleus, and RA, the primary motor nucleus. RA projects directly onto brainstem nuclei with outputs to respiratory and syringeal muscles. Learning pathway: LMAN and Area X in the forebrain and DLM in the midbrain are active (and required) during song development. Cytochrome oxidase stained image courtesy of S. Helekar.



Auditory

Mld = midbrain nucleus receives afferents from medulla – analogous to mammalian inferior colliculus

Ov = auditory thalamic nucleus

Field L = primary auditory area in telencephalon

NCM = secondary auditory area, projections to HVC shelf

CM = secondary auditory area, projections to HVC shelf

Nif = sensorimotor nucleus, afferents from auditory areas, primary auditory input to HVC

HVC = premotor song nucleus (proper name)

Motor and Learning

Nif = auditory input to song system

HVC = premotor song nucleus (proper name)

RA = motor nucleus, outputs to brainstem motor nuclei

Area X = song nucleus involved in learning, afferents from HVC

LMAN = song learning nucleus with direct projections to primary motor nucleus of the song system, RA

DLM = thalamic nucleus, receives projections from Area X sends projects to LMAN

1.4. Sensory tuning in the zebra finch

Is there direct evidence that sensory and perceptual changes might be involved in song learning?

For example, can we see in the motor song system, evidence for changes in tuning with regards to tutored song, or to BOS, as it develops? Perceptual tuning, changes in behavioral response to sounds,

has not been measured systematically in the zebra finch over development. Studies of sensory tuning, changes in physiological responses to sounds over development, indicate developmental sensory tuning in song control and song learning nuclei. Sensory tuning is commonly measured by single or multi-unit responses to a battery of stimuli: Bird's Own Song, Tutor Song (BOS), Conspecific song (CON), and heterospecific song. To test specificity of tuning to spectral and temporal features of song stimuli may be manipulated, for example, BOS played in reverse maintains spectral features but distorts the temporal order of these features. Using these methods developmental sensory tuning has been shown in the song system nuclei, including the learning pathway. The song system premotor nucleus (HVC) exhibits dynamic auditory responses in relation to BOS and tutored song, with stronger neural responses for TUT early in sensorimotor learning and increased responses to BOS as the bird learns his song (Nick and Konishi, 2005). In the learning pathway as well, Area X and LMAN nuclei are increasingly responsive to BOS, as the bird is practicing his song (Solis and Doupe, 1997). Sensory tuning in the song motor and learning pathways correlate with behavior of the animal: neurons become tuned to BOS when the bird is actively practicing his song (while presumably matching it to the auditory template). After hearing and memorizing the tutor song, the most relevant auditory signal for the bird's output is his own current vocalization.

What about sensory tuning to song in the auditory system? Does it precede sensory tuning in the song system? Primary auditory region Field L develops in complexity and selectivity in birds that have a rich auditory environment between the sensory and sensorimotor stages of song development (Gehr Hofer, Marquadt, Leppelsack, 2000). But increased selectivity is primarily due to emergence of subregions that deal with complex tasks (e.g. code the frequency modulation of syllables) and individual neurons in this region do not develop song-selectivity between day 30-60 days (Gehr, et al., 2000). In young birds (35 days old) neural responses in Field L and CLM are also generally depressed and highly variable (compared to adults) and do not show song-selectivity (Amin, et al., 2007). As early as 20 days post-hatch secondary auditory area NCM has robust responses to song stimuli and preference (in some neurons) for CON over heterospecific song (Stripling, Kruse, Clayton, 2001). In sum, developmental sensory tuning towards biologically relevant sounds - BOS, the vocalization the bird is working on – is seen in song motor and learning pathways. But sensory tuning over development in the auditory system is not as clear: responses of single neurons are highly variable in the young bird, and do not show clear developmental changes in preference for the Bird's Own vocalizations.

1.5. Integration

Comparing findings on the sensory side of vocal learning across humans and songbirds, we have seen that in human infants it is not possible to separate innate components from experience because even the earliest exposure to spoken language shapes auditory responses and perception in a complicated manner. In songbirds, we can control the early auditory environment and show conclusively that some of the sensory filters, or biases for song (conspecific over heterospecific) are indeed innate, and with a controlled environment we could see how experience influences sensory tuning.

1.6. Rationale for Methods

Methods that are commonly used in humans to study development of auditory sensitivity and perception – ERP, fMRI, behavioral tasks have not been widely used, or until now, not available for use in songbirds (review, Van der Linden, Van Meir, Boumans, Poirier, Balthazart, 2009). In this study we aimed to implement ERP and fMRI methods in zebra finches. These methods probe auditory sensitivity as well as perception across the temporal and spatial domains and have been used extensively in humans to identify patterns of global brain responses that correspond to auditory processing of non-speech sounds, individual speech sounds, speech contrasts, and groups of speech sounds (language) (Kuhl and Rivera-Gaxiola, 2008).

Functional Magnetic Resonance Imaging uses changes in magnetic field to measure the haemodynamic response related to neural activity in the brain. The Blood Oxygen-Level Dependent (BOLD) signal can be localized and parametric maps provide a visual representation of the spread and intensity of activity in response to stimuli or cognitive tasks.

Auditory event related potentials use far-field (scalp) electrodes to track stimulus induced time-locked changes in electrical activity in the auditory pathway and cortical areas (Key, Dove, Maguire, 2005; Luck, 2005). These responses lie on a spatial and temporal spectrum: immediate (0-15ms) responses generated by the cochlea, auditory nerve, and auditory brainstem used to assess peripheral and brainstem auditory function; mid-latency responses (50-150 ms) presumably reflect initial activation of auditory cortex; and long-latency responses, or Cortical Auditory Evoked Potentials (CAEP's) – 150ms – 300ms and longer) - generated by multiple sources in primary and association auditory cortices reflecting low- and high-level auditory processing (Luck, 2005). The evoked response is time-locked and reflects synchronous firing of neural populations along the auditory pathway; the signal provides good temporal resolution of the 'global' (i.e. organism-level) auditory response. Together the spatial resolution of the fMRI signal and temporal resolution of the ERP signal are complementary.

Studying songbirds has one great advantage over studying infants: the environment of the bird can be controlled throughout the entire development of the bird. The auditory environment of the birds can be tightly controlled: birds can be raised in acoustic isolation until they can be operantly trained to express preference (as juveniles); this can test innate auditory biases. Exposure to song can be controlled, in order to test the limits of song learning and the effect of auditory experience on production as well as perception.

Here we use noninvasive functional Magnetic Resonance Imaging (fMRI) and ERPs to localize and determine time course of song and syllable driven brain activity. The general aim is to find neural representations of song at the global level and determine a quantitative measure of auditory perception in the vocal learning zebra finch. Quantifying the global representation of song and song syllables in the adult brain will lead us to the question of how that representation forms in the juvenile during song development. In order to establish a foundation for performing comparative studies of auditory processing related to vocal development in humans and zebra finches we set the following aims:

Aim 1a: Implement the non-invasive brain imaging technique, functional Magnetic Resonance Imaging, in zebra finches.

Functional Magnetic Resonance Imaging (fMRI) has been used extensively in humans to identify patterns of global brain responses that correspond to individual speech sounds, speech contrasts, and groups of speech sounds (language) (Kuhl and Rivera-Gaxiola, 2008). In the songbird, the animal model of human speech development, the most common research method is an invasive and terminal electrophysiological preparation. Single and multi-unit electrophysiology has been widely used to describe neural circuitry and fine-grained neural responses of sensory and motor nuclei in the song system (Solis and Doupe, 1997; Woolley and Casseday, 2004). How can such different physiological responses be compared? And how can these responses be compared across species? We propose using non-invasive fMRI in the zebra finch to study spatial brain topography of sensory development in vocal learning. In order to implement this technique in zebra finches and compare responses to those in humans a number of technical and theoretical issues must be addressed:

- a) Brain size, composition and complexity is extremely different between the two species, but fMRI hardware and image resolution is the same, what adjustments can and must be made to improve spatial resolution in the much smaller brain?
- b) The explosion of fMRI studies in humans has resulted in well-established Regions of Interest (ROI) and anatomical landmarks for automatic registration of functional activity onto anatomical maps. Using this technology in zebra finches requires exploration in terms of stimuli, areas and amplitude of activity, and mapping functional activity onto anatomical plots (without well-established landmarks);
- c) Humans are cognizant and willing subjects, generally able to follow directions for scanning. Laboratory animals may be stressed in experimental preparations; stressed zebra finches are extremely active and attempt to fly. How can we minimize stress of the bird (and consequent movement and movement artifacts in image)? Will the bird require a restraint, sedation or anesthetization?
- d) Using a common brain imaging technique in humans and zebra finches will provide some agreement between the research findings in these two species, i.e. similar spatial scales of response relative to the respective brain. But comparing actual brain responses of the two species in a meaningful way is uncharted territory.

In order to address these issues we will explore parameters of scan design and procedure:

- Determine number of stimulus repetitions required to elicit activation and obtain adequate spatial resolution (i.e. scan time versus image resolution).
- Determine general Region of Interest for auditory activation in the zebra finch brain.
- Determine type and duration of stimulus that will elicit activation.

- Determine minimal adequate restraint and sedation (if necessary) to minimize stress and produce usable images.

Aim 1b: Implement the physiological technique, Event Related Potentials, in zebra finches.

Auditory Event Related Potentials (ERPs) have been widely used in humans to study maturation of the auditory system (Ponton, Eggermont, Kwong, Don, 2000) experience-dependent auditory processing of speech sounds (Dehaene-Lambertz, Dehaene, Hertz-Pannier, 2002). Experimental electrophysiology in songbirds generally focuses on firing patterns of a small number of individual or multiple neurons in specific brain nuclei across birds. There have been limited attempts to map auditory responses of neurons throughout an entire nucleus (HVC) (George, Cousillas, Richard, Hausberger, 2003); and a number of studies have used auditory evoked responses as an index of hearing sensitivity and function in the periphery and brainstem (Dooling, 1975; Brittan-Powell and Dooling., 2004). Far-field ERP recordings should be a relatively easy technique to implement in the zebra finch (compared to micro-electrode single unit recordings) to record global brain activity. But this technique has been overlooked because responses are global rather than localized. In human literature it is recognized that global neural activity may provide neural correlates of perception (Rodriguez, George, Lachaux, Martinerie, Renault, Varela, 1999), and bridge molecular and cellular activity with behavioral responses. We propose to develop the ERP technique in zebra finches in order to have a basis for comparing auditory neural responses to biologically relevant vocalizations across the two species. Using Event-Related Potentials would provide the same relative spatial scale of response (global activity) in the two species and these responses can be compared in common parameters: frequency bands of activity, latency and amplitudes of responses. Developing this technique in zebra finches will require us to:

- a) Determine practical level of invasiveness.** In protecting the brain, the skull also diffuses electrical signals from internal sources complicating source localization. In humans, scalp electrodes record this diffused signal through the skull. The skull-effect on the ERP signal in zebra finches is currently not known: the adult zebra finch skull is double-layered with an air-filled space between the two layers; given this construction, neural signal might be impossible to detect on the outside of the skull. With epidural electrodes we can measure global brain response and remove one obstacle to obtaining the signal.
- b) Determine parameters for obtaining ERP signal in zebra finches:** What kind of material to use for chronic electrodes and how many electrodes to use? What stimuli elicit evoked response? How many stimulus presentations are required to obtain measurable evoked response? How can we minimize stress and movement in the bird during Evoked Potential recordings? How to analyze and compare evoked potential recordings between electrodes, stimuli, and birds.
- c) Examine potential signal sources.** As in the fMRI technique, as of yet, there are no established generators or components of the ERP signal in the zebra finch. We will explore

responses to different stimuli and observe response patterns and remark on potential signal generators.

Hypothesis II : Reliable and interpretable fMRI BOLD responses can be obtained from adult zebra finches in a 3.0 Tesla magnet imaging machine using an auditory paradigm.

Hypothesis I : Reliable and interpretable auditory event-related potential data can be obtained from epidural electrodes in adult zebra finches.

Aim 2: Using fMRI and ERP examine the role of early sensory versus early sensory-motor experience in shaping auditory responses to vocal sounds.

Establish techniques and then control sensory environment during development and ask questions about how manipulation of sensory environment affects auditory responses in adult birds. We would like to ask how vocalizations and auditory response changes in relation to each other during developmental vocal learning, but because vocal learning is dynamic we expect auditory responses also to be changing over development. In order to judge developmental responses we need to obtain data in (presumably) more stable adults. In adult birds we can still ask questions about development by controlling the developmental experience of birds and looking at the outcome in adult responses.. Conventional wisdom tells us that experiences shapes individuals but how can the effect of experience, specifically sensory experience, be measured or detected in the individual? And can common experiences of individuals be detected in a group effect? We ask whether the ERP and fMRI responses will allow us to detect the kind of auditory and song learning experiences birds have had during development. We will do this by comparing responses across groups of birds that have had different auditory, social, and song learning experiences during development.

Hypothesis I : Developmental auditory experience qualitatively affects adult auditory responses to natural sounds in male zebra finch, who learn song during development, as measured by Event Related Potentials and functional Magnetic Resonance Imaging.

Hypothesis II : Minimal auditory experience with a song model during development is sufficient to shape auditory responses of adult male birds.

Hypothesis III : Lack of developmental auditory experience (external song template) will qualitatively and quantitatively affect auditory responses in the adult male, as it affects vocal production in the adult male.

Hypothesis VI : Developmental auditory experience differentially shapes auditory responses in male and female zebra finches.