Cover Page





The handle http://hdl.handle.net/1887/31633 holds various files of this Leiden University dissertation.

**Author**: Kant, Anne Marie van der
**Title**: Neural correlates of vocal learning in songbirds and humans : cross-species fMRI studies into individual differences
**Issue Date**: 2015-01-28

# Neural correlates of vocal learning
# in songbirds and humans

cross-species fMRI studies
into individual differences

Anne M. van der Kant

Anne M. van der Kant

Neural correlates of vocal learning in songbirds and humans:
cross-species fMRI studies into individual differences

PhD Thesis, Leiden University & University of Antwerp

Faculteit der Farmaceutische, Biomedische
en Diergeneeskundige Wetenschappen

# Neural correlates of vocal learning
# in songbirds and humans

cross-species fMRI studies
into individual differences

## Neurale correlaten van vocaal leren
## bij zangvogels en mensen

fMRI studies naar soortelijke en individuele verschillen

(met een samenvatting in het Nederlands)

## Proefschrift

voorgelegd tot het behalen van de graad van
doctor in de Biomedische Wetenschappen
aan de Universiteit Antwerpen
te verdedigen door

## Anne Marie van der Kant

# Neural correlates of vocal learning
# in songbirds and humans

cross-species fMRI studies
into individual differences

## Proefschrift

## Anne Marie van der Kant

geboren 4 mei 1984
te Lelystad

Never lose your sense of wonder

# Contents

CHAPTER 1

## General introduction

*Animal models, songbirds particularly, are increasingly used to study the human capacity for speech and language. In the light of understanding both language evolution and individual language acquisition these models are highly valuable, provided that they are studied within a valid comparative framework. In the past few decades, non-invasive methods such as functional Magnetic Resonance Imaging (fMRI) and Near-InfraRed Spectroscopy (NIRS) have become available for human as well as animal brain research. In the studies discussed in this thesis, fMRI is employed to unravel the neural correlates of vocal learning in the human and songbird brain. The rationale for this route towards a better understanding of the human language capacity will be outlined in this introduction, exploring both (1) how cross-species comparative research, especially comparisons between humans and songbirds, can inform questions about human language acquisition and (2) how recently developed non-invasive neuro-imaging techniques can contribute to our understanding of the neural substrates of vocal learning in general and the human capacity for language in particular.*

# 1.1 What can birdsong learning teach us about human language?

## 1.1.1 The value of the songbird model for human speech

Human language is widely regarded as the pinnacle of the uniqueness of human cognition within the animal kingdom. This is a valid standpoint in so far that only humans show evidence of the ability to develop and use a communication system showing the full complexity of human language. However, other animal species do possess the ability for vocal communication and vocal learning, although we have to search outside the primate lineage to find species with these abilities (Egnor & Hauser, 2004). Songbirds share the ability to learn their communicative signals through vocal imitation with humans and a small number of other species (e.g. hummingbirds (Baptista & Schuchmann, 1990), parrots (Pepperberg, 2002), cetaceans (Reiss & McCowan, 1993), pinnipeds (Ralls, Fiorelli, & Gish, 1985; Sanvito, Galimberti, & Miller, 2007), bats (Janik & Slater, 1997) and possibly elephants (Poole, Tyack, Stoeger-Horwath, & Watwood, 2005)). With approximately four thousand species, oscine songbirds represent the largest group of vocal learners in the animal kingdom. These songbirds learn to produce their song by imitating the vocalizations of their male adult conspecifics.

Most species that show vocal learning differ greatly from humans in evolutionary descent, brain and behavior. The limited vocal learning abilities in primates, our closest relatives, complicate the study of vocal learning from the perspective of common descent. Moreover, birdsong, which is the most studied model for human language, is highly limited in its productivity and does not show symbolic representation or duality of patterning like human language (Hockett, 1963). Furthermore, human language is characterized by its capacity to express and combine meaning in an unrestricted way. In contrast, birdsong can have different functions, including mate attraction and territorial defense, which are not coupled to the acoustic content of the song. Although songbirds have been claimed to show some combinatorial potential (e.g. Gentner, Fenn, Margoliash, & Nusbaum, 2006), these findings are controversial (Stobbe, Westphal-Fitch, Aust, & Fitch, 2012; Van Heijningen, De Visser, Zuidema, & Ten Cate, 2009). These factors limit the possibilities for the extrapolation of findings from birdsong studies to the study of human language.

However, birdsong learning, especially in the zebra finch *(Taeniopygia guttata)*, does show some striking parallels with the way in which human infants acquire their native language. More specifically, several aspects of the zebra finch song learning process, among others the memorization of the target song during the sensory phase (Eales, 1985) and motor practice or 'babbling' during the sensorimotor phase (Marler, 1970), are comparable to processes which were observed in human speech learning (Doupe & Kuhl, 1999). Furthermore, zebra finches as well as humans are hypothesized to be age-limited learners,

implying that they both need to be exposed to adult vocalization within a specific sensitive period during development in order to develop normal song or language. By focusing on these characteristics of development, songbird studies can add to the understanding of neural mechanisms underlying the development of speech in human infants and of vocal learning in humans as a species. Many other aspects of language, like the capacity for symbolic reference, can be found in other lineages like parrots (e.g. Pepperberg, 2002) and non-human primates (Pollick & De Waal, 2007).

Although non-human primates do not show vocal imitation learning, they do exhibit comprehension learning and usage learning (e.g. alarm calls in vervet monkeys (Seyfarth, Cheney, & Marler, 1980)). The study of apes and monkeys can add to the understanding of human language development in several ways. They can serve as a non-vocal imitation learning control species when comparing vocal learning in humans and birds and additionally they can provide insights into other behaviors that might have contributed to human language evolution, such as auditory processing (Petkov, Kayser, Augath, & Logothetis, 2006), gesture (Pollick & De Waal, 2007), joint attention (Carpenter & Tomasello, 1995) and symbolic reference (Ribeiro, Loula, De Araújo, Gudwin, & Queiroz, 2007). Because chimpanzees and bonobos are our closest relatives, some of the general learning mechanisms they possess may share their evolutionary origin with learning mechanisms that were adapted to sub-serve language in human evolution. By comparing humans to both songbirds and non-human primates, mechanisms we might have in common with either vocal learners or non-vocal learners can be identified.

Studies of animal analogues of capacities needed for human language can complement each other in order to establish which shared cognitive mechanisms underlie the human capacity for language. Because animal studies allow for manipulation of different environmental and developmental factors, both behavioral and neural mechanism can be studied in isolation, which is very difficult to accomplish in human studies. By studying behavioral mechanisms that have been shown to be comparable between species as opposed to attempting to compare complete communication systems, birdsong research has the potential to uncover mechanisms underlying human language that are highly challenging to study in humans.

However, because cognitive mechanisms underlying vocal behavior and auditory processing in songbirds and humans originate from highly dissimilar avian and mammalian brains, comparative studies should take brain function into consideration. Possible approaches that take this issue into consideration are discussed in the second section of this introduction. The studies discussed in the present thesis aspire to develop an understanding of the similarities and differences between the neural processing of learned species-specific vocalizations in avian and mammalian vocal learners.

Subsong / Plastic song    Crystallization

Memorization

DPH  0      20      40      60      80     100     120
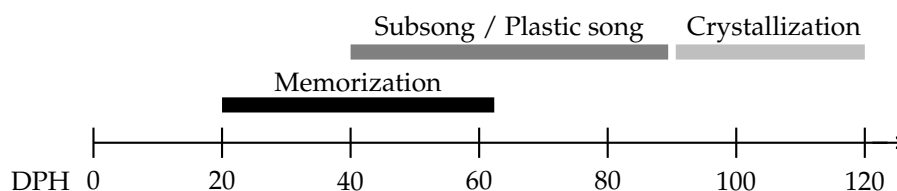
Figure 1.1: Stages of song learning in zebra finches. Memorization of the tutor song starts before juveniles start producing subsong. During the sensorimotor phase juveniles produce subsong and plastic song and use auditory feedback in order to match their vocalizations to the memory of the tutor's song. At the end of the song learning process, one song is crystallized and will be sung by the bird for the rest of its life. DPH = Days Post Hatching.

### 1.1.2   Learning by imitation in songbirds and humans

Songbirds as well as human infants acquire their species-specific communicative signals by vocal imitation learning. This process involves (i) a preference for species-specific and in particular the parents' vocalizations from a very early developmental stage, (ii) memorization of the vocalizations of adult conspecifics, (iii) motor practice where the juvenile bird or human infant practices sounds that were memorized and (iv) the use of auditory feedback in order to compare one's own vocalizations with the memorized adult vocalizations. Evidence for each of these elements has been found in songbirds as well as in humans.

In order to learn the speech or song of their adult conspecifics, both human infants and juvenile songbirds need to know which sounds to imitate. Evidence for such an innate preference for species-specific vocalizations has indeed been found. Human babies prefer the language spoken by their mother from very early on (Moon, Cooper, & Fifer, 1993). Although macaques do not imitate vocalizations, some auditory regions in the macaque temporal lobe are selectively activated by conspecific sounds and even by the voice of individual conspecifics (Petkov et al., 2008). These results suggest that some brain functions that are important in vocal learning were in place before vocal imitation developed in the human lineage. Zebra finches show a preference for conspecific vocalizations even when reared in isolation (Braaten & Reynolds, 1999). Furthermore, auditory fMRI studies in zebra finches, have shown selectivity for conspecific song and the song of individual birds in the auditory midbrain nucleus, the avian homologue of the inferior colliculus (Poirier, Boumans, Verhoye, Balthazart, & Van der Linden, 2009; Van der Kant, Derégnaucourt, Gahr, Van der Linden, & Poirier, 2013) (see also Chapter 3).

Studies in human infants as well as in juvenile zebra finches suggest that they not only prefer, but also memorize adult vocalizations. Newborns can discriminate familiar stories that were read to them in the womb from unfamiliar stories (DeCasper & Spence, 1986), while the zebra finch brain shows selectivity for the tutor song (usually the father's), even in adulthood (see

Chapter 3 or Van der Kant et al. (2013)). Further evidence for a memory of the tutor's song comes from the fact that male zebra finches are able to produce a highly accurate copy of the adult song they were exposed to during the sensory phase of song learning (Figure 1.1), even if the song model is no longer present during vocal motor practice (Funabiki & Konishi, 2003). However, for the memorization of adult vocalizations, purely the presence of conspecific speech or song is not sufficient. Social interaction is highly important for vocal learning in both humans and songbirds. Human infants benefit from joint attention during social interactions with adults for the development of their vocabulary (Tomasello & Farrar, 1986). The same holds true for zebra finches, for which live tutoring or tutoring with a plastic model bird (Beecher & Burt, 2004) leads to better imitation compared to tape tutoring (Ten Cate, 1991) or operant tutoring (for a comparison of different methods, see Derégnaucourt, Poirier, Van der Kant, Van der Linden, & Gahr, 2013). This suggests that attention plays a large role in the ability to memorize adult vocalizations.

Both the production of birdsong and human speech start with a phase where repetitive strings of unstructured vocalizations are produced that increasingly resemble adult speech sounds or song notes. In songbirds this is known as subsong (Marler, 1970), which is produced early in the sensorimotor phase of song learning (Figure 1.1), while it is referred to as babbling in human infants. Babbles or subsongs are produced independently of a specific social context (Lenneberg, Rebelsky, & Nichols, 1965) and develop around the same developmental stage in both normal hearing and hearing-impaired infants (Stoel-Gammon & Otomo, 1986). Babbling thus appears to be a form of vocal motor practice that develops spontaneously and is only later influenced by social interaction (Goldstein, King, & West, 2003) and auditory feedback. Babbling behavior appears to be common among vocal learning species as it was also shown in bats (Knörnschild, Behr, & Von Helversen, 2006).

During babbling and subsequent phases of song and speech learning, auditory feedback plays an important role in the development of adult vocalizations. Auditory feedback is needed in order to evaluate whether one's own vocalizations match the memory of the tutor song (in birds) or adult speech sounds (in babies). Both in songbirds and human infants, the absence of auditory feedback impairs vocal learning. Congenitally deaf babies initially start to vocalize, but fail to develop more structured vocalizations with true consonants due to the absence of auditory feedback (Stoel-Gammon & Otomo, 1986). Furthermore, song development is heavily compromised if songbirds are denied auditory feedback after sensory learning (Konishi, 1965), despite the presence of a fully formed memory of the tutor song. Studies on acquired deafness in humans (Cowie, Douglas-Cowie, & Kerr, 1982) and deafening (Nordeen & Nordeen, 1992) as well as pertubation of auditory feedback (Leonardo & Konishi, 1999) in songbirds have demonstrated that auditory feedback is not only crucial for song and speech learning, but also for its subsequent maintenance.

### 1.1.3   Sensitive periods for birdsong learning and human language acquisition

As discussed in the previous sections, songbird studies addressing vocal learning behavior have revealed a number of striking similarities with human language acquisition, which have contributed to our understanding of learned vocalizations. An important parallel on which this thesis focusses is the existence of a sensitive period for vocal learning in both humans (Lenneberg, Chomsky, & Marx, 1967) and some songbirds (Eales, 1985). This "sensitive period hypothesis" is supported by a number of findings from both birdsong and human speech.

For human speech, support for the sensitive period hypothesis is predominantly based on individual cases. Theoretically, the hypothesis can be tested by withholding a child spoken language from birth and ask whether this child develops spoken language on its own. Because of the cruelty in such an isolation experiment, it cannot be carried out in humans and isolation studies are limited to case studies of children who grow up isolated from human contact. Recent studies addressing the language capacities of such a "wild" child have shown that absence of human language input impedes he development of spoken language and especially normal syntax (e.g Curtiss, Fromkin, Krashen, Rigler, & Rigler, 1974; Fromkin, Krashen, Curtiss, Rigler, & Rigler, 1974). Moreover, the spoken language abilities of prelingually deaf children who receive a cochlear are related to the age at implantation (Nikolopoulos, O'Donoghue, & Archbold, 1999). Another indication for a sensitive period for language acquisition with a high level of plasticity early in life comes from the finding that infants perceive small differences between both native and non-native speech sounds, while perception becomes increasingly categorical with a tendency to place non-native speech sound in native categories during the first year of life (Werker & Tees, 1983). Furthermore, second language acquisition is known to proceed with less effort and more successfully at younger ages (e.g. Johnson & Newport, 1989), though the loss of vocal learning ability appears to be more gradual than previously assumed (Hakuta, Bialystok, & Wiley, 2003).

Songbird species differ considerably in the timing of their song learning. Seasonal learners, like starlings (Eens, Pinxten, & Verheyen, 1992), exhibit a period of plasticity at the start of each breeding season, where they start to sing more often and add new songs to their repertoire, while open-ended learners, like canaries (Nottebohm & Nottebohm, 1978), are able to continuously learn new songs throughout their lives. Zebra finches, however, show a pattern that is more similar to human spoken language development, where they learn their song early in life and must be exposed to adult vocalizations within a restricted time window of approximately 120 days, the sensitive period for song learning, in order to develop normal song (Eales, 1985). When reared in isolation of adult song, zebra finches develop an abnormal "isolate" song (Thorpe, 1958). There are some instances in which the song can change,

including deterioration due to deafness (Nordeen & Nordeen, 1992) or adaptation to distorted auditory feedback (Leonardo & Konishi, 1999). Furthermore, a prolongation of the sensitive period was observed in birds that showed incomplete learning at the end of the period after which zebra finch song is usually fully developed (Derégnaucourt et al., 2013).

The behavioral phenotypes of the sensitive periods for vocal learning in songbirds and humans show some interesting similarities. Furthermore, sensitive periods are not unique to vocal learning, but are widespread during the development of sensory systems, among which binocular vision (e.g. Mower, 1991) is the best known example. This raises the question whether the nature and timing of these sensitive periods is governed by domain-general neural mechanisms, which may be shared between species.

### 1.1.4   Different brains, similar mechanisms?

Both vocal imitation learning and the presence of a sensitive period for vocal learning are shared between songbirds and humans. These similarities suggest that both species have an innate predisposition for vocal learning, possibly originating from shared neural mechanisms. However, large differences in brain anatomy pose challenges to comparative brain research in songbirds and humans. When conducting comparative research, addressing the neural mechanisms underlying vocal learning in songbirds and humans, the large structural differences between the mammalian and avian brain have to be taken into consideration. The bird brain does not have a cortex and is structured in nuclei rather than lamina. Moreover, different views exist as to how the bird brain and the mammalian brain evolved from a common ancestor, implicating different homologies between brain structures (Jarvis et al., 2005). The different structures of the zebra finch brain that are implicated in song learning and production are further illustrated in Figure 1.2 and the regions supporting language functions in the human brain are depicted in Figure 1.3.

However, despite the lack of consensus about the common descent of specific neural structures, functional analogies can inform theories about convergence on the neural level which might underlie the parallels in vocal learning behavior. For example, the brain of vocal learning songbirds as well as the human brain show neural pathways for the processing and production of learned vocalizations, which are not found in avian and primate species without vocal learning capacities (e.g. Jarvis, 2004; Petkov & Jarvis, 2012). Furthermore, recent studies have shown similarities in gene expression between songbird and human brains in regions related to vocal learning (e.g. Pfenning et al., 2014). For example, Hara, Rivas, Ward, Okanoya, and Jarvis (2012) compared levels of parvalbumin (PV), a calcium binding protein in vocal learning and non-vocal learning birds and found significant differences in PV expression between distantly related avian vocal learners and non-vocal learning species. These findings suggest that the neural architecture for vocalization in species which show vocal learning is fundamentally different from the neural
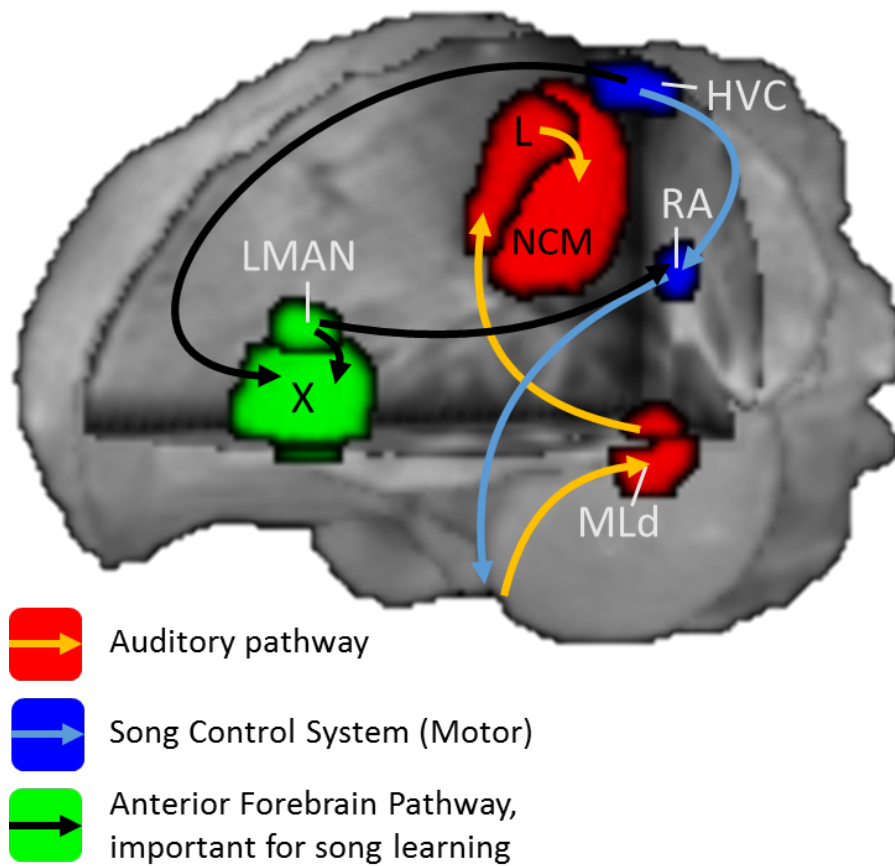
**Figure 1.2: Songbird brain regions implicated in singing, song processing and song learning.**
MLd: Lateral Mesencephalic Nucleus, avian homologue of ICC, L: Field L, primary auditory area, NCM: Caudomedial Nidopallium, secondary auditory area, HVC: Song production planning, RA: nucleus Robustus of the Arcopallium, song motor nucleus, X: Area X, LMAN: Lateral Magnocellular nucleus of the Anterior Nidopallium. Area X and LMAN are part of the Anterior Forebrain Pathway, the pathway linked to song learning.

architecture of non-vocal learning species, even if they are phylogenetically close.

Moreover, the songbird brain exhibits a specialized forebrain circuit, which plays a crucial role in birdsong learning (Brainard & Doupe, 2000; Sohrabji, Nordeen, & Nordeen, 1990) and possibly subserves the generation of subsong (Aronov, Andalman, & Fee, 2008), but is not necessary for adult song production (Nordeen & Nordeen, 1993). This circuit shows similarities with the human basal ganglia, which have recently been found to be functionally connected to regions important for language processing (Booth, Wood, Lu, Houk, & Bitan, 2007) and to be implicated in syntactic processing (Kotz, Frisch, Von Cramon, & Friederici, 2003).

Adult production of birdsong and possibly speech are subserved by neural networks distinct from those that subserve vocal learning (see Figures 1.2 and 1.3). In contrast to juvenile babbling or subsong, adult song production allows for monitoring of vocalizations, but is limited in its experience-dependent plasticity (Brainard, 2004). This mechanism is supported by the finding that adult monkeys, birds and humans all show suppression of auditory neurons during own vocalizations (Eliades & Wang, 2008; Houde, Nagarajan, Sekihara, & Merzenich, 2002; Müller-Preuss & Ploog, 1981; Numminen, Salmelin, & Hari, 1999). The involvement of brain regions that allow for a reduced experience dependent plasticity of vocalizations might underlie sensitive periods for vocal learning in both songbirds and humans.

Furthermore, the genome of the zebra finch has recently been sequenced (Warren et al., 2010), showing striking similarities between these birds and humans with regard to genes involved in vocal communication. This work sets the stage for further studies into the genetic basis of vocal learning, showing the involvement of FoxP2, in zebra finch song learning (e.g. Haesler et al., 2007; Teramitsu, Poopatanapong, Torrisi, & White, 2010). Defects in FoxP2 have previously been associated with an inherited speech disorder in humans (Lai, Fisher, Hurst, Vargha-Khadem, & Monaco, 2001).

To conclude, although the brain of a songbird differs significantly from ours and does not show the same architecture as a mammalian brain, similar neural mechanisms may drive the development, perception and production of learned vocalizations. Although the abovementioned parallels on the neural level all provide pieces to the puzzle, overarching theories, addressing the common neural basis of vocal learning, are still lacking. This thesis aims to contribute to the understanding of common neural mechanisms underlying vocal learning in songbirds and humans. In order to compare the neural mechanisms underlying both learning and perception of species-specific vocalizations, this thesis presents three functional Magnetic Resonance Imaging studies, addressing the neural substrate of (i) perception of learned vocalizations in adult songbirds (Chapter 3), (ii) birdsong learning within the sensitive period for vocal learning (Chapter 4) and (iii) human artificial grammar learning and perception of the learned grammar (Chapter 5). Additionally,
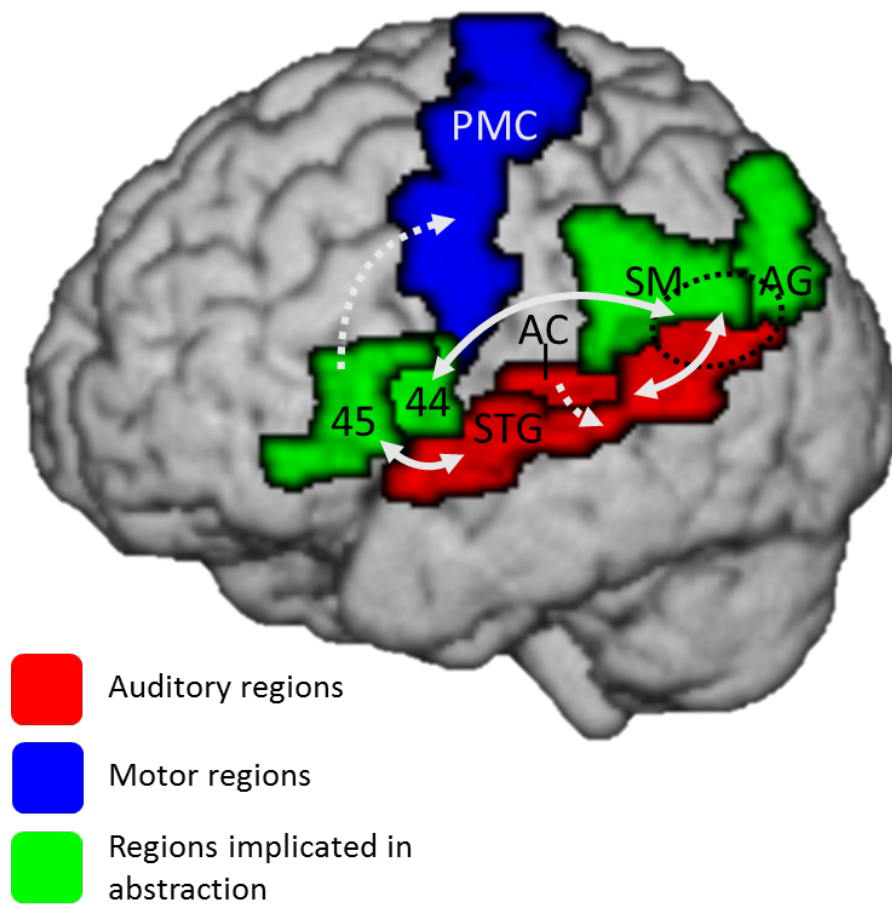
**Figure 1.3: Human brain regions implicated in language processing.** AC: Auditory Cortex, STG: Superior Temporal Gyrus, PMC: Primary Motor Cortex, 45: BA 45 or Pars Triangularis of the Inferior Frontal Gyrus, 44: BA 44 or Pars Opercularis of the Inferior Frontal Gyrus (BA 44 and 45 are part of Broca's area, which is implicated in syntactic processing), AG: Angular Gyrus, SM: Supramarginal Gyrus. The dotted oval shows the location of Wernicke's area, which is implicated in semantic processing.

the contribution of brain connectivity to learning new linguistic structures in addressed in Chapter 6.

## 1.2 fMRI for comparative studies on vocal learning

### 1.2.1 Methods for songbird and human brain research

Both on the behavioral and the neural level, parallels between birdsong and human language have been demonstrated (for a review, see Doupe & Kuhl, 1999). However, studying birds and humans often requires different methodologies which have traditionally been used in separate fields. Therefore, a large number of parallels have been drawn based on studies conducted in a single species. Recently, song and speech learning have been compared using similar behavioral methods (Gobes et al., 2009; Ter Haar, 2013), but no comparative studies addressing the neural correlates of birdsong learning and human language acquisition have been conducted to date.

Traditionally, studies addressing the neural substrates underlying birdsong learning have employed invasive methods such as electrophysiology and IEG expression (for a review see Bolhuis & Gahr, 2006). These types of studies have contributed greatly to our understanding of the neural substrates of birdsong learning and the song system. Moreover, they were the basis for a number of hypotheses about the cellular and molecular basis of human language acquisition (Doupe & Kuhl, 1999). However, these methods might not be best suited for comparative studies involving humans, because they focus on the cellular level, while the non-invasive methods typically applied in human research operate on the neuronal population level.

For example, electrophysiological research in non-human primates has led to the discovery of mirror neurons (Rizzolatti, Fadiga, Gallese, & Fogassi, 1996). Songbirds have also been shown to possess auditory forebrain neurons that are both activated during song production and listening to the same song (Prather, Peters, Nowicki, & Mooney, 2008). However, these studies used single-cell measurements, which cannot be employed in human participants, because this method is highly invasive. Human fMRI studies, on the other hand, cannot establish whether specific cells have true mirror neuron properties, although the hypothesized right-hemisphere human mirror neuron system shows striking similarities with left-hemisphere networks involved in human language processing and production (Iacoboni, 2005). In this case, songbirds can inform questions about the role of mirror neurons in vocal learning and the use of songbird (and primate) fMRI in combination with electrophysiology can aid the validation of human fMRI results.

Songbird studies aiming to understand vocal learning processes in humans search for common neural or genetic mechanisms to explain common behavior, but direct comparisons between the neural substrates of human language and birdsong development have yet to be made. In order to achieve

a sound comparison between species, ideally, the same experimental methods are used. Because studies of human brain function typically involve non-invasive (neuro-imaging) methods such as EEG or (functional) MRI, the Bio-Imaging Lab in Antwerp has taken the approach to implement fMRI in zebra finches (Boumans, Theunissen, Poirier, & Van der Linden, 2007). Because fMRI has also been established in non-human primates (Logothetis, Guggenberger, Peled, & Pauls, 1999), the neural correlates of vocal learning and memory can now be effectively compared between songbirds, non-human primates and humans. The studies described in the present thesis employ auditory fMRI in zebra finches and human subjects in order to uncover the neural substrates of auditory vocal learning in both species.

### 1.2.2   fMRI in songbirds and humans

Due to its non-invasiveness, using fMRI to study the neural correlates of birdsong learning has several advantages. Firstly, we can compare the development of the neural substrates for vocal learning in songbirds and humans. Secondly, because animals are not sacrificed after the experiment, longitudinal studies are possible and finally, the whole brain can be studied, which facilitates the study of large-scale brain changes. This technique thus enables us to longitudinally follow birdsong development in the same individuals and to consequently compare these data to fMRI data on human language development or auditory perception and memory in non-human primates.

In recent years, auditory fMRI in zebra finches has revealed highly specific responses to species-specific sounds (Boumans, Vignal, et al., 2008), the bird's own song (Poirier et al., 2009) and the song of the bird's father, which was learned during development (see Chapter 3 or Van der Kant et al. (2013)). Furthermore, the technique to acquire auditory fMRI data from juvenile zebra finches has been developed and applied for the first time for the research reported in the present thesis (see Chapter 4). These developments, as well as the technical details and challenges of songbird fMRI, are reviewed in Chapter 2.

Although the results are promising, there are still some limitations to inter-species comparative studies using (f)MRI. One of these limitations is the need for anesthesia when collecting fMRI data from small songbirds. Like humans, non-human primates like macaques (Petkov et al., 2006, 2008) as well as pigeons (De Groof et al., 2013) can be measured in an awake state, although a significant amount of training is needed. However, fMRI in awake zebra finches has yet to be established and the skittish nature of these small songbirds might complicate training. Furthermore, the extensive training periods required for awake animal fMRI experiments limit the possibilities for developmental studies. Anesthesia might also influence the BOLD response, which complicates the comparison with humans. Specifically, isoflurane, which is used for zebra finch fMRI studies described in this thesis, works through the GABA-ergic system and might thus influence responses in GABA-ergic neu-

rons. Indeed, in a number of brain regions, selective responses to the bird's own song appear to depend on cognitive state (Cardin & Schmidt, 2004; Dave, Albert, & Margoliash, 1998). Measuring songbirds under anesthesia also means that stimulation of the auditory system is not influenced by attention, which is the case in human fMRI studies.

Although anxiety is less of an issue in human subjects, including children (Westra et al., 2011), in The Netherlands, inclusion of children younger than eight years in MRI experiments without anesthetic in a non-clinical setting is not allowed (Central Committee on Research Involving Human Subjects, 2004). Therefore, fMRI experiments in children can only be conducted after the start of first language acquisition. Near InfraRed Spectroscopy (NIRS), like fMRI, is based upon a hemodynamic signal. This method is suitable for use in infants, even newborns (Gervain, Macagno, Cogoi, Peña, & Mehler, 2008), because the signal is obtained by measuring infrared light, scattered by the cortex, directly on the scalp. Because light is used and surces and detectors are placed directly on the scalp, this method does not involve a confined space, a magnetic field or radio-frequency pulses. However, the light does not penetrate very deep, limiting measurements to the outer layers of the cortex.

Not only the need for anesthesia, but also the nature of birdsong and human language pose challenges to comparative studies. The complexity of human language requires the isolation of a single aspect of language learning or speech perception in order to test specific predictions about this process. On the other hand, tasks can and need to be more challenging for humans and participants can be asked for behavioral responses, which allows for the testing of more intricate processes like grammar learning. Zebra finches only learn a single song, implying that all that has been learned can be tested in a single experiment. However, due to the anesthesia, behavioral responses cannot be obtained and the need for a social context excludes the possibility of tutoring a songbird in the scanner and observing the learning process.

These practical differences between human and songbird fMRI studies complicate the comparison between the BOLD responses. However, if the limitations are taken into consideration, human and songbird fMRI both provide information about the haemodynamic response to auditory stimulation with species-specific vocalizations that can be either learned or non-learned. Because the level at which the measurements are taken (larger groups of neurons) and the neural mechanism that is measured (the BOLD response) are the same, songbird and human fMRI studies provide a manner through which hypotheses regarding neural processes underlying vocal learning, stated based on previous behavioral and invasive studies, can be assessed using a comparative approach.

In the studies described in the present thesis, experimental conditions could not be kept exactly the same for human subjects and songbirds. In contrast to songbirds, human subjects cannot be isolated from species-specific vocalizations and thus will never enter into an fMRI experiment naive. Conse-

quently, studying the acquisition of elements from a known natural language is problematic.

In order to overcome the language experience that adult humans necessarily have, we employed artificial grammar learning to study language acquisition from mere exposure in adult humans. An artificial grammar was introduced to serve as a new language-like stimulus that could be learned in the course of an experiment and consequently tested. While the learning process and its outcomes in songbirds are studied in a juvenile and an adult group, in humans the process is simulated in a single group of adult subjects.

## 1.3  Research questions and hypothesis

In this introduction, the framework was outlined within which this thesis employs fMRI in both the songbird model and human artificial grammar learning to study the neural underpinnings of human language acquisition. Furthermore, some of the methodological challenges faced in comparative research studying both human and animal brains have been discussed. In the studies described in the experimental part of this thesis, fMRI in both songbirds and humans will be used to study the neural mechanisms underlying birdsong learning and human artificial grammar learning. In a series of fMRI studies investigating these neural mechanisms in adult and juvenile zebra finches and human adults, this thesis compares the neural substrates of song learning in birds with those of language learning in humans. With these studies, it aims to shed light on the following research questions:

1. To what extent do the neural substrates of birdsong learning and human language acquisition show similarities?

2. Can potential similarities be ascribed to common neurocognitive mechanisms underlying the development of both birdsong and human language?

3. Are our fMRI results able to strengthen the birdsong model with a brain-based account of the behavioral similarities found in songbirds and humans with regard to vocal learning?

We hypothesize that in both species, vocal learning will "tune" the brain towards a state where it shows selective responses to learned species-specific communication signals (e.g. a specific (artificial) grammar in humans or a specific song in birds). Therefore, a song or a grammar that was perceived before, specifically if it was learned within the sensitive period for song or language development, will elicit a larger activation in regions that are implicated in learning species-specific vocalizations. The development of selectivity for learned vocalizations might represent a common neural mechanism underlying vocal learning.

## 1.4   Contributions

Chapter 1: conception, literature study and paper and manuscript writing: Anne van der Kant

Chapter 2: conception: Annemie Van der Linden, paper writing: Lisbeth Van Ruijssevelt and Anne van der Kant, manuscript supervision: Geert De Groof and Annemie van der Linden

Chapter 3: conception and experimental design: Colline Poirier, bird rearing: Sébastien Derégnaucourt, data collection: Anne van der Kant, data analysis and paper writing: Anne van der Kant and Colline Poirier, manuscript supervision: Annemie Van der Linden and Manfred Gahr

Chapter 4: conception and experimental design: Anne van der Kant, bird rearing and data collection: Anne van der Kant, data analysis and paper writing: Anne van der Kant and Lisbeth Van Ruijssevelt

Chapter 5: conception and experimental design: Anne van der Kant, data collection: Anne van der Kant and Daan van de Velde, data analysis and manuscript writing: Anne van der Kant, manuscript supervision: Niels Schiller and Claartje Levelt

Chapter 6: conception and experimental design: Anne van der Kant, data collection: Anne van der Kant and Daan van de Velde, data analysis and manuscript writing: Anne van der Kant, manuscript supervision: Niels Schiller

Chapter 7: conception, literature study and manuscript writing: Anne van der Kant

# Part I

# Neural correlates of vocal learning in zebra finches

In the first chapters of this thesis, the songbird model will be employed in order to study the neural correlates of vocal learning in a controlled environment. There are several reasons why songbirds provide a more controlled environment in which to study these neural correlates. Firstly, human infants start to learn language even before they are born and can not be isolated from any of their language experience, early or later, for obvious ethical reasons. Furthermore, human language is a complex system with many types of rules, an extensive vocabulary and phonological system. Therefore, the acquisition of all the knowledge a human infant needs to speak and understand language cannot be studied in a single experiment. Zebra finch males, on the other hand, learn a single song of which the perception can be studied in isolation and compared with other songs.

Taking this approach, the following chapters explore how the zebra finch brain processes the song that is learned from the tutor in comparison to other songs and shed light on the developmental trajectory of tutor song processing. In order to visualize the processing and memorization of the learned song in the tiny zebra finch brain, auditory fMRI is used. The methodology of the application of this technique in songbirds is discussed in Chapter 2. In order to compare the neural correlates of birdsong learning and human language learning, in part II of this thesis, language learning in human adults is simulated using an artificial grammar learning paradigm and the learning process is studied using the same auditory fMRI method.

CHAPTER 2

# Current State-of-the-Art of auditory functional MRI (fMRI) on Zebra Finches: Technique and Scientific Achievements

*Songbirds provide an excellent model system exhibiting vocal learning associated with an extreme brain plasticity linked to quantifiable behavioural changes. This animal model has thus far been intensively studied using electrophysiological, histological and molecular mapping techniques. However, these approaches do not provide a global view of the brain and/or do not allow repeated measures, which are necessary to establish correlations between alterations in neural substrate and behaviour. In contrast, Functional Magnetic Resonance Imaging (fMRI) is a non-invasive in vivo technique which allows one (i) to study brain function in the same subject over time, and (ii) to address the entire brain at once. During the last decades, fMRI has become one of the most popular neuroimaging techniques in cognitive neuroscience for the study of brain activity during various tasks ranging from simple sensorimotor to highly cognitive tasks. By alternating various stimulation periods with resting periods during scanning, resting and task-specific regional brain activity can be determined with this technique. Despite its obvious benefits, fMRI has, until now, only been sparsely used to study cognition in non-human species such as songbirds. The Bio-Imaging lab (University of Antwerp, Belgium) was the first to implement blood oxygen level dependent (BOLD) fMRI in songbirds, and in particular zebra finches, for the visualization of sound perception and processing in auditory and song con-*

*trol brain regions. The present chapter provides an overview of the establishment and optimization of this technique in the Bio-Imaging laboratory and of the resulting scientific findings, some of which are discussed more elaborately in Chapters 3 and 4. The introduction of fMRI in songbirds has opened new research avenues that permit experimental analysis of complex sensorimotor and cognitive processes underlying vocal communication in the songbird model.*

## 2.1 Introduction

Over the past decades, the songbird model has become widely used in neuroscientific studies addressing changes in brain function and structure. The main reason for the great interest in this animal model is that songbirds exhibit vocal learning associated with extreme brain plasticity linked to quantifiable behavioral changes. Consequently, these animals present currently the best available model to study neuronal plasticity accompanying normal daily processes such as learning but also recovery from neurological insults. Zebra finches, for example, are widely used to study the neuroprotective role of sex steroids following traumatic brain injury in different brain regions (e.g. Duncan & Saldanha, 2011; Mirzatoni, Spence, Naranjo, Saldanha, & Schlinger, 2010; Peterson et al., 2007; Walters, Alexiades, & Saldanha, 2011). In addition, an international team of scientists has recently decoded the genome of the zebra finch (Warren et al., 2010). This study revealed striking similarities between these birds and humans with regard to genes involved in vocal communication. This work thus sets the stage for future studies that could help identify the genetic and molecular origins of speech disorders, such as those related to autism, stroke, stuttering and Parkinson's disease.

In order to further unravel the mechanisms of vocal learning and related disorders, integrated brain studies in the songbird model should simultaneously consider functional and morphological changes and cellular function. This is only possible by using in vivo techniques based on a whole-brain approach allowing the visualization of structural brain changes, interactions between brain regions and/or the activity in defined neuronal populations. Magnetic resonance imaging (MRI) allows repeated non-invasive measures of both brain structure and activity. By visualizing the whole brain over the course of longitudinal studies, the time frame and spatial coordinates of crucial events can be determined for further in depth invasive molecular analysis of the underlying mechanisms. Different MRI sequences can now be applied to animal species as small as mice or songbirds. Diffusion tensor imaging, for example, has been implemented in starlings and can be used to visualize fiber tracts and their changes (De Groof & Van der Linden, 2010). In addition, resting state functional MRI was successfully used in mice (Jonckers, Van Audekerke, De Visscher, Van der Linden, & Verhoye, 2011) and could in the future be used in songbirds to determine functional connectivity between brain regions. Finally, the Bio-Imaging Lab also implemented functional MRI (fMRI)

in songbirds as small as zebra finches for the visualization of sound perception and processing (e.g. Boumans, Gobes, et al., 2008; Boumans et al., 2007; Boumans, Vignal, et al., 2008; Poirier et al., 2011; Van der Kant et al., 2013).

The present review provides an overview of the establishment and optimization of auditory BOLD fMRI in songbirds and the subsequent scientific findings that have permitted the visualization of sound perception and processing in auditory and song control brain regions of songbirds. The introduction of fMRI in songbirds opened new research avenues that allowed for the experimental analysis of complex sensorimotor and cognitive processes underlying vocal communication in songbirds and has begun to be implemented by others (Maul et al., 2010; Voss, 2011; Voss, Salgado-Commissariat, & Helekar, 2010; Voss et al., 2007).

### 2.1.1 Functional Magnetic Resonance Imaging (fMRI)
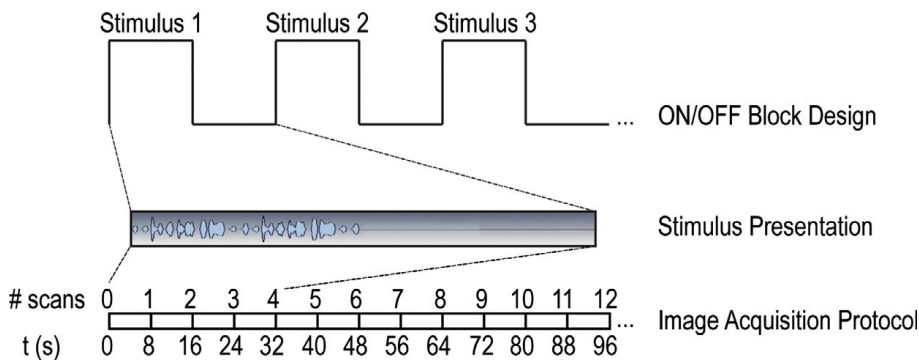


**Figure 2.1: BOLD fMRI data acquisition.** Schematic representation of a typical ON/OFF auditory stimulation block design in which auditory stimulation periods are alternated with rest periods. In this example, each block (stimulus/rest) lasts 16 s during which 2 MR images are acquired. The different stimuli consist of representative motifs of birdsong or other types of sound depending on the experiment. Within 1 fMRI experiment, the stimuli of interest (typically 23 per experiment) are repeated several times in a pseudo-random order with the same number of presentations for each stimulus.

fMRI is a powerful tool for studying in vivo brain function in humans and animals. The fMRI technique, which is most frequently used to image brain activity, relies on Blood Oxygenation Level Dependent (BOLD) contrast (Ogawa et al., 1992). BOLD fMRI detects the global hemodynamic changes in response to (synaptic and/or electrical) activity of a large number of neurons. This technique allows repeated analysis of information processing by neuronal networks during various tasks ranging from simple sensorimotor to highly cognitive tasks. Although fMRI has been increasingly used in human cognitive studies, its use in small animal models has been limited due to the required immobilization and anaesthesia during imaging, which lim-

its the type of questions that can be addressed (Van der Linden, Van Camp, Ramos-Cabrer, & Hoehn, 2007).

The BOLD contrast reects a local change in the ratio between oxygenated and deoxygenated haemoglobin in the blood induced by a local increase in oxygen consumption during neuronal activation. Deoxygenation of haemoglobin causes its iron to become paramagnetic, which inuences the magnetic field experienced by the protons in the surrounding water molecules (Ogawa, Lee, Kay, & Tank, 1990). During neural activity, an increase in oxygen consumption is followed within a few seconds by a larger fractional increase in blood ow and an increase in blood volume, resulting in a net decrease in the amount of deoxygenated haemoglobin (Malonek et al., 1997). In this case, the capillary and venous deoxyhaemoglobin concentrations decreases, leading to a decreased distortion of the magnetic field which will in turn increase the MRI time constant $T_2^*$. This increase is reected as an elevation of signal intensity in $T_2^*$-weighted magnetic resonance images. Advanced image processing can detect the local changes in the magnetic field in a series of images acquired during and/or after stimuli presented in a typical block design (Figure 2.1). In this way, resting and task-specific regional brain activity can be calculated.

### 2.1.2 Auditory fMRI of speech and vocal communication

In humans, communication through learned spoken language provides an opportunity to design auditory stimulation paradigms for fMRI ranging from simple sound stimulation (Stevens & Weaver, 2009) to word or sentence recognition (Meltzer, Postman-Caucheteux, McArdle, & Braun, 2009), memorization (Mano et al., 2011) and grammar learning (Chapters 5 and 6. A few animal groups (songbirds, hummingbirds and parrots and possibly whales, dolphins, and bats) have developed a learned vocalization system used in social interactions. Extensive neuro-ethological research in songbirds has yielded significant insight into how their brain generates and perceives species-specific vocalizations in different social contexts and developmental stages. Although vocal communication presumably emerged independently in songbirds and humans (Hauser, Chomsky, & Fitch, 2002), intriguing parallels between birdsong and human speech have been identified (for reviews see Doupe & Kuhl, 1999; Kuhl, 2003; Wilbrecht & Nottebohm, 2003), including the proposed existence of sensitive periods for learning, dependence of this process on auditory experience and feedback, and lateralization of sound production (e.g. Cynx, Williams, & Nottebohm, 1992; George, Cousillas, Richard, & Hausberger, 2005; Phan & Vicario, 2010).

Interestingly, the songbird's auditory and sensorimotor brain regions involved in song acquisition and production show reliable electrophysiological responses to acoustic stimuli even under anaesthesia (Capsius & Leppelsack, 1996; Grace, Amin, Singh, & Theunissen, 2003; Schmidt & Konishi, 1998). In contrast to the human brain, the songbird brain can also easily be experimentally manipulated. This model system is therefore ideally suited for fMRI

studies and provides a unique model for studying the complex cognitive processes underlying vocal communication. The song control and auditory brain regions of songbirds are well documented and intensively studied using both electrophysiological and activity-dependent immediate early gene (IEG) techniques. These studies have shown that song control nuclei respond selectively to the bird's own song (BOS) (Margoliash, 1986) and that secondary auditory areas show a specialization in conspecific song processing (Grace et al., 2003; Mello, 2004).
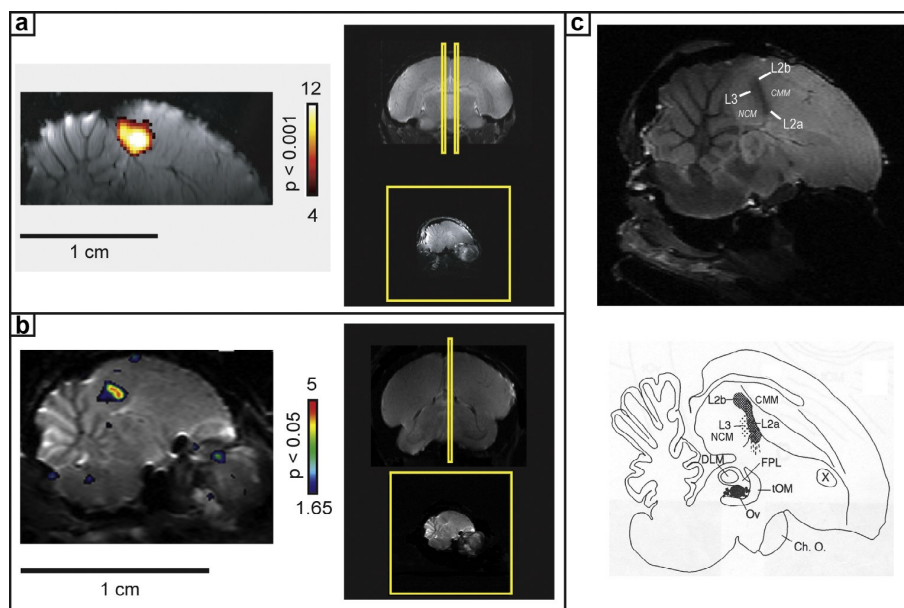


**Figure 2.2: First assessment of brain function in songbirds by auditory BOLD fMRI.** (a and b) Z-score maps illustrating brain activation as observed by BOLD fMRI in response to conspecific song in male starlings (a) and zebra finches (b). Z-values are color coded according to the scale displayed in the figure and only voxels in which the t-test was found to be significant are displayed. The MRI images were acquired according to the slice positioning presented in the middle. Modified from data in Boumans, Vignal, et al. (2008) and Van Meir et al. (2005). (c) High resolution $T_2$-weighted Spin Echo (SE) MR image of the male zebra finch auditory system illustrating Field L (top). The schematic representation (bottom) illustrates how the dense fibertract which defines subfield L2a, corresponds to the darker ellipsoid in the MR image. Ch.O., optic chiasm; CMM, caudomedial mesopallium; DLM, medial nucleus of the dorsolateral thalamus; FPL, lateral forebrain bundle; L2a, L2b, and L3, subregions of Field L; NCM, caudomedial nidopallium; Ov, nucleus ovoidalis; tOM, tractus occipitomesencephalicus; X, area X. Redrawn from figures in Boumans, Vignal, et al. (2008) and Vates et al. (1996).

However, activity-dependent expression of IEG and electrophysiological recordings of single- or multi-unit activity are highly invasive or even lethal methods. In addition, electrophysiological techniques are limited by the number of locations that can be sampled in one experiment and thus require a priori hypotheses about the localization of the neuronal substrate involved in the

process investigated. In contrast, BOLD fMRI is non-invasive, allows a whole-brain approach and can be used to perform assumption-free experiments.

The use of BOLD fMRI in songbirds to study the processing of auditory stimuli has been initiated and developed at the Bio-Imaging Lab (University of Antwerp, Belgium) and further applied in collaboration with a substantial list of songbird neuroscientists (e.g. Boumans et al., 2007; Boumans, Vignal, et al., 2008; Van Meir et al., 2005). The technique is currently being exploited to address cognitive research questions in songbirds (De Groof, Poirier, George, Hausberger, & Van der Linden, 2011; Poirier et al., 2009; Poirier, Verhoye, Boumans, & Van der Linden, 2010; Van der Kant & Van der Linden, 2012). In addition, auditory fMRI has been used to map the auditory regions of the brain (Van Meir et al., 2005), investigate spectral and temporal processing in the auditory forebrain (Boumans, Gobes, et al., 2008; Boumans et al., 2007; Boumans, Vignal, et al., 2008) and identify the neural substrates responsible for the recognition of the BOS (Poirier et al., 2009).

## 2.2    BOLD response in the auditory system of a robust songbird: feasibility and technical challenges

### 2.2.1    Implementation of BOLD fMRI in the starling brain

The feasibility of auditory fMRI in songbirds was first demonstrated in 2005 in starlings (Van Meir et al., 2005) before being extended to a smaller songbird species, the zebra finch (e.g. Boumans et al., 2007) (Figure 2.2). The first fMRI study in (song)birds established the existence of a hemodynamic response and demonstrated that the temporal pattern of the auditory BOLD response is remarkably similar to auditory evoked BOLD responses in human subjects (Van Meir et al., 2005). Although the BOLD signal mainly reects local field potentials and thus differs from the action potentials measured in electrophysiological and IEG studies (e.g. Logothetis, Pauls, Augath, Trinath, & Oeltermann, 2001), BOLD fMRI has confirmed many results obtained by these techniques. Our first fMRI study in starlings demonstrated the successful application of complex auditory stimulation in the noisy environment of the MRI-scanner. By exposing the birds to white noise, music and conspecific song, we could confirm the specialization of the caudal telencephalon for processing conspecific songs over artificial sounds.

### 2.2.2    Technical challenges

The use of BOLD fMRI in songbirds poses practical problems at the level of brain imaging, stimulus presentation, and data processing. We will briey summarize the technical issues involved in detecting the BOLD signal in the brain
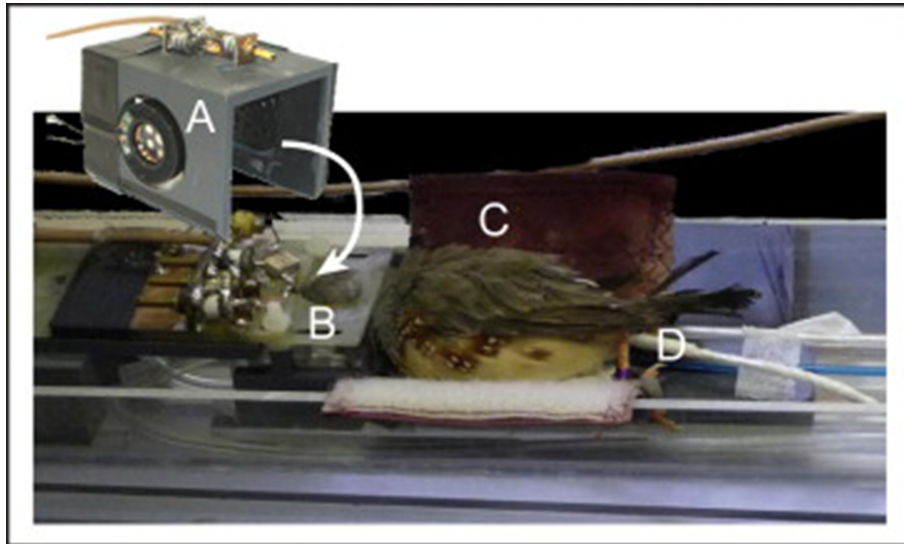
**Figure 2.3: Setup for auditory fMRI in zebra finches.** (A) Helmholtz RF transmission antenna to which magnetless dynamic speakers for the presentation of the auditory stimuli, are fixed; (B) custom-made head holder composed of a beak mask and a circular radio-frequency surface antenna tightly placed around the bird's head to allow accurate reproducible positioning of the bird while at the same time preventing motion; (C) restraining jacket; (D) cloacal temperature probe connected to a feedback controlled heater system to keep the body temperature of the bird stable during the measurement. In addition, a pneumatic pillow sensor is placed under the bird to monitor the respiration rate. Redrawn from figures in Poirier et al. (2011) and Van Meir et al. (2005).

of a songbird during auditory stimulation together with a description of the solutions adopted.

Earlier MRI of songbird brains produced high resolution images by application of $T_1$- and $T_2$-weighted Spin Echo (SE) sequences at 7T (Van der Linden et al., 1998; Van Meir et al., 2004; Verhoye et al., 1998). However, imaging with classical $T_2^*$-weighted Gradient Echo (GE) sequences that are mostly used to detect hemodynamic susceptibility changes in BOLD fMRI is affected in birds due to severe susceptibility artefacts generated by the air cavities that developed in the skull.

Auditory stimulation in the noisy environment of the scanner due to the switching gradients is also problematic. Uncoupling the auditory stimulus from image acquisition often solves this problem in human research because the maximal amplitude of the BOLD response is only reached 310 s after stimulus onset. Fast GE planar images with very long repetition times are often acquired using a sparse imaging paradigm to visualize activated regions (for reviews see Amaro et al., 2002; Moelker & Pattynama, 2003). Since temporal aspects and optimal stimulus duration to obtain maximal BOLD responses in songbirds were unknown at this stage, this approach was not an option. Based

on a careful evaluation of available data, it was decided to start BOLD fMRI studies in songbirds with a GE-Fast Low Angle Shot (FLASH) sequence. GE-FLASH offers a spatial resolution that is high enough to visualize small telencephalic auditory regions located in brain regions suffering the least from susceptibility artefacts induced by air cavities in the skull. In order to maximize the temporal resolution of the GE FLASH fMRI acquisition, only 1-2 slices per animal were imaged (Figure 2.2). The slice positioning was determined depending on the specific questions addressed in the experiment. In order to reduce the gradient's noise, we used a long gradient ramp time (1000 ls), producing a sustained humming background noise of 7080 dB, during both stimulation and rest periods. Birds were exposed to acoustic stimuli with a maximum sound pressure level of 100 dB as commonly used during auditory fMRI studies in humans (Jäncke, Shah, Posse, Grosse-Ryuken, & Müller-Gärtner, 1998). The head was positioned in a custom made stereotaxic device, leaving the ears free for auditory stimulation by a non-magnetic dynamic speaker (Figure 2.3).

Another important issue in performing songbird fMRI relates to the anesthetized condition of the subjects during imaging. Most studies in humans are carried out in an awake state, whereas small animals need to be immobilized during imaging. In consequence, when fMRI is performed in animals, they need to be anesthetized or curarized (Peeters, Tindemans, De Schutter, & Van der Linden, 2001) or trained to tolerate restraining in a noisy scanner environment (Ferris et al., 2005, 2008; Liang, King, & Zhang, 2011; Sachdev et al., 2003; Zhang et al., 2010). The latter is time consuming and hence only few experiments have been carried out with this approach. Moreover, since it is difficult for awake restrained animals to focus on a specific task, this procedure potentially increases inter-individual variation, which could mask activation patterns and stimulus specificity (Peeters et al., 2001). In our studies, we therefore decided to visualize the activity of auditory regions in anesthetized subjects. This approach is supported by previous studies on humans showing by a combination of EEG and fMRI that the sleeping brain can still process auditory stimuli and detect meaningful acoustic events (Portas et al., 2000) and is further supported by the above mentioned responses to acoustic stimuli of the song control and auditory regions under anaesthesia in songbirds (Capsius & Leppelsack, 1996; Grace et al., 2003; Schmidt & Konishi, 1998).

## 2.3 BOLD fMRI in the zebra finch's auditory system

After the development of the technique in a large songbird, the starling, auditory BOLD fMRI was implemented in zebra finches (Figure 2.2). This small Australian songbird is a widely used animal model in many biological disciplines. The great advantage of using zebra finches over other songbirds is

that they are commercially available, easy to breed in captivity and robust to anaesthesia. They currently present one of the best available and most frequently used animal models to study brain changes during speech development. Zebra finches are age-limited learners. Like humans, they learn their vocalizations early in life during a specific sensitive period for vocal learning. During the first part of their sensitive period, the sensory phase, zebra finch male chicks develop a model based on the tutor song they hear. Subsequently, they use this model to develop their own song during vocal practice, which starts during the sensorimotor phase of song learning. By the end of the sensorimotor phase (around 80-100 days post hatching (DPH)) every individual male has learned to sing one particular song with a relatively simple structure (Brainard & Doupe, 2002). Each male's song is different, although birds of the same bloodline will exhibit similarities, and all finches will overlay their own uniqueness onto a common rhythmic framework. The song of a zebra finch lasts for only a few seconds and has a simple syntactic and temporal organization. A typical adult male zebra finch song is preceded by a series of introductory notes. Syllables are the simple elements, separated by silence, which in turn consist of notes, the boundaries of which are defined by abrupt changes in acoustic characteristics. A song motif or phrase is made up of a string of syllables lasting between 0.5 and 1.5 s. The motif is repeated multiple times in a song, and the song can be repeated several times to form a bout (Sossinka & Böhner, 1980).

For the implementation of in vivo BOLD fMRI in zebra finches, various preliminary experiments were performed in our laboratory in order to explore the possibility to visualize auditory processing in the auditory brain regions of these small songbirds. Subsequently, we investigated neural activity in response to different auditory stimuli as well as the effect of different orders of song manipulation. (Boumans, Gobes, et al., 2008; Boumans et al., 2007; Boumans, Vignal, et al., 2008). In an initial study, the behavioral response to songs played inside and outside the magnet bore was shown to be similar (Boumans, Vignal, et al., 2008). In order to validate our results, the BOLD response measured with fMRI was confirmed by Near-infrared Spectroscopy (NIRS) (Vignal et al., 2008) and with IEG expression study (Boumans, Vignal, et al., 2008; Vignal, Attia, Mathevon, & Beauchaud, 2004).

### 2.3.1 Visualization of the BOLD response in the zebra finch's auditory system: validation with in vivo optical methods

Simultaneously with the application of BOLD fMRI to zebra finch brain activity during auditory processing, the efficiency of the method for visualizing physiological hemodynamic brain responses was tested in a hypercapnia model (Vignal et al., 2008). Hypercapnia or hyperoxia is widely used as a global physiological stimulus to alter the BOLD contrast (Jones, Berwick, Hewson-Stoate, Gias, & Mayhew, 2005; Martin, Jones, Martindale, & Mayhew,
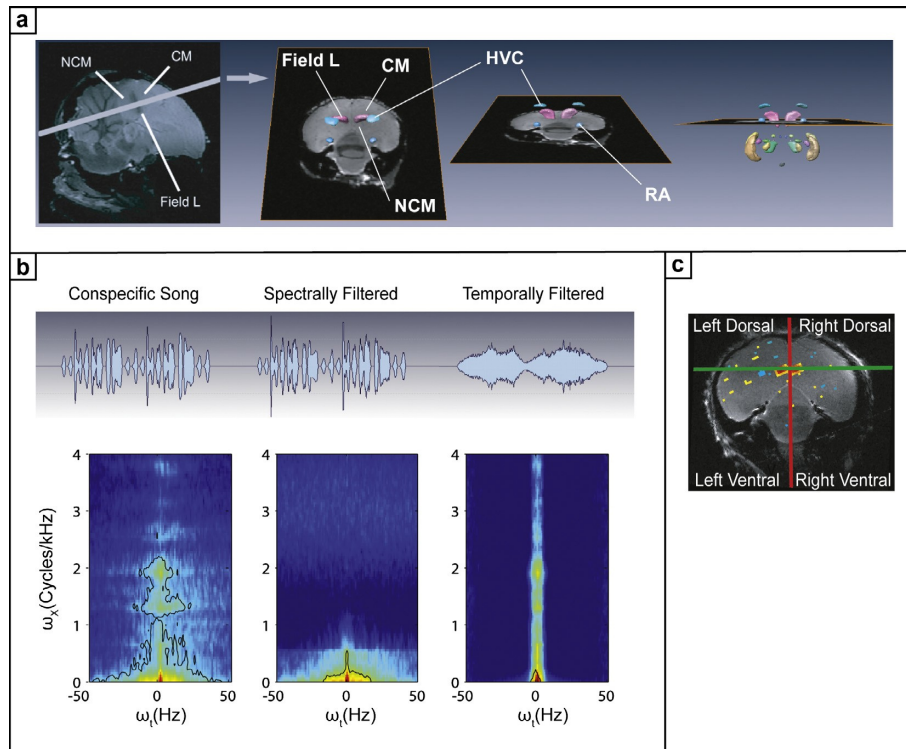
**Figure 2.4: Imaging auditory processing of the spectral and temporal features of song in the auditory forebrain of zebra finches.** (a) Slice positioning. To discern the neural substrates of spectral and temporal features of song in the zebra finch, a tilted coronalhorizontal slice which comprised most of the auditory regions in the zebra finch, was analyzed. This slice passed through a relatively anterior region of the caudal mesopallium (CM), a central region of Field L (mainly L2a), a caudal region of the caudomedial nidopallium (NCM) and song control robust nucleus of the arcopallium (RA). The imaging slice is displayed in a 3D rendering volume of the zebra finch brain atlas representing some of the key auditory regions and song control nuclei (right) (Poirier et al., 2008). (b) Oscillograms (top) and modulation spectra (bottom) showing the spectral and temporal features found in the experimental auditory stimuli. The figure shows an example of conspecific song before and after spectral and temporal filtering. $\omega_\chi$ = spectral modulations, $\omega_t$ = temporal modulations. (c) Regional analysis. The auditory regions are divided into four quadrants determined by the midline dividing the right from left hemisphere, and by the horizontal fiber tract dividing ventral from dorsal regions. The dorsal quadrants include the dorsal part of L2a, L1 and L. L2b, while the ventral quadrants include the ventral part of L2a, L3, the medial-ventral part of L and NCM. Redrawn from figures in Boumans et al. (2007).

2006; Schwarzbauer, Mildner, Heinke, Brett, & Deichmann, 2010). It induces a well known vasodilatation response resulting in an increase in cerebral blood ow (CBF) associated with a small rise in cerebral blood volume (CBV) without change in oxygen consumption (Dutka, Scanley, Does, & Gore, 2002; Lee, Duong, Yang, Iadecola, & Kim, 2001; Wu, Luo, Li, Zhao, & Li, 2002). As expected, a significant increase in BOLD signal during the hypercapnic period was detected using fMRI. These findings were validated using Near Infrared Spectroscopy (NIRS). The results showed a good correlation between the local haemodynamic changes visualized by BOLD fMRI and the variations in total haemoglobin and oxygen saturation level measured using NIRS.

## 2.3.2 BOLD response in zebra finch's auditory system upon robust $1^{st}$ order manipulations of conspecific song stimuli: validation with immediate early gene expression and behavioral responses

In addition to optical methods, an IEG expression study was used to further validate the technique. In parallel to a study performed by the same collaborating research group using IEG ZENK expression as a measure of brain activity in awake birds (Vignal et al., 2004), the Bio-Imaging Lab studied the discriminatory properties of auditory forebrain regions using BOLD fMRI in anesthetized male zebra finches (Boumans, Vignal, et al., 2008). The results from the ZENK expression study demonstrated the ability of the zebra finch brain to discriminate biologically relevant information (conspecific song) when mixed with different levels of broadband noise. In addition, behavioral responses to conspecific song as well as IEG expression in the secondary auditory region NCM (caudomedial nidopallium) showed a progressive decrease in response to increasing background noise similar to the fMRI study. These findings suggest that NCM plays a role in the ability to discern song from masking noise. The fMRI study confirmed these results in anesthetized animals and thus shows that the neural responses in the auditory system of anesthetized and awake zebra finches to the stimuli applied in this study are highly comparable. Additionally, the BOLD fMRI study revealed neural activity in Field L, the primary auditory center, elicited by conspecific song which could not be shown by IEG expression (Mello, Vicario, & Clayton, 1992). In contrast to the response in NCM, this BOLD response is not reduced by added noise.

For this fMRI study, the preservation of the behavioral response to the song stimuli played during the experiments was tested to control for potential deformations of the sound in the restricted environment of the scanner bore. The stimuli were recorded within the bore and played to awake and freely behaving zebra finches followed by a behavioral analysis. Results showed that the number of distance calls emitted by the experimental birds in response to conspecific song recorded in the bore did not differ from the amount of calls emitted in response to songs that were recorded outside the scanner environment

(Boumans, Vignal, et al., 2008).

### 2.3.3 BOLD response in the zebra finch auditory system upon $2^{nd}$ order spectral and temporal song stimuli manipulations

The discovery of NCM as a primary auditory region involved in segregation of conspecific song from background noise raised new questions about the way vocalizations are processed in the avian brain. In a new fMRI study using synthetic song in which only the temporal or only the spectral structure of natural conspecific song was preserved (Figure 2.4b), we investigated which centers in the avian auditory forebrain play a role in the processing of the spectral and temporal structure found in the succession of song-syllables (Boumans et al., 2007). In addition, we investigated the neural recognition of the BOS in brain regions involved in song production and learning (the song nuclei) as this requires the preservation of both the spectral and temporal structure of song (Theunissen & Doupe, 1998). In the course of the implementation of BOLD fMRI in the zebra finch, this study was the first quantitative assessment of the distribution of the auditory evoked BOLD response in the zebra finchs forebrain. At this stage, the spatial resolution of the structural MR-images allowed only a rough delineation of the different auditory regions. As the principal auditory evoked BOLD response to the presented stimuli (BOS, multiple familiar conspecific songs and spectrally filtered or temporally filtered versions of these conspecific songs) was in Field L and centered around L2a (Figure 2.2c), a regional analysis was performed by dividing the auditory regions into four quadrants: the right and left hemisphere, and the ventral and dorsal regions of Field L. Based on prior measurements and the comparison to earlier published anatomy, the dorsal quadrants included the dorsal part of L2a and L1, L and potentially L2b, while the ventral quadrants include the ventral part of L2a, L3 and, in some experiments, the medial-ventral part of sub-region L as well as potentially NCM (Figure 2.4a and 2.4c). The ventral regions of Field L including subarea L3, medial-ventral subarea L and potentially NCM showed depressed responses to spectrally filtered song. Field L as a whole and especially the dorsal regions showed increased responses to temporally filtered song which represented more song stimuli per time unit in this study. This study revealed no differences between BOS and conspecific song across the analyzed regions.

### 2.3.4 BOLD response in the zebra finch's auditory system upon $2^{nd}$ order song manipulations and the bird's own song

Within the previous study mentioned above (Section 2.3.3, Boumans et al. (2007)), regions with preferential responsiveness to the BOS compared to other

familiar conspecific songs, were addressed. On that account, a tilted coronal-horizontal slice was analyzed in order to sample most of the auditory regions in the zebra finch forebrain. This slice passed through a relatively anterior region of the caudal mesopallium (CM), a central region of Field L (mainly L2a) and a caudal region of NCM (Figure 2.4a and 2.4c). In this set of scans, no significant BOS selectivity or hemispheric differences could be discovered. In a next study, we continued our search as we believed that determining which regions show preferential responsiveness to the BOS is of great importance. Neurons sensitive to self-generated vocalizations could mediate the auditory feedback process necessary for the detection of errors in the bird's own vocal production which can subsequently be corrected to stabilize song. In order to visualize BOS selective regions, the next dataset was acquired in the orthogonal direction, i.e. the sagittal plane (Boumans, Gobes, et al., 2008). This plane passed through the regions L2a, L3 and NCM, as well as caudal L2b and the caudal region of the medial CM (CMM) that were not investigated in our previous study. The goal of this new study was twofold. First, we measured the global neural activity elicited by the BOS and conspecific song in sub-regions L2a and L2b. Because of its anatomical connections, we suspected that the sub-region L2b might be involved in an intermediary processing step between L2a and the secondary auditory regions, and that it could thus be the original locus from where BOS selectivity emerges. Secondly, we examined the general selectivity for characteristic acoustical features found in song over synthetic sounds as it was previously observed at the single neuron level (Amin, Doupe, & Theunissen, 2007; Grace et al., 2003; Hsu, Woolley, Fremouw, & Theunissen, 2004). In contrast to the previous study (Boumans et al., 2007) in which only $2^{nd}$ order song manipulations were performed, we compared responses to normal BOS with synthetic variations on the BOS that differed in spectro-temporal and/or modulation phase structure. These so called $3^{rd}$ order manipulations preserved the overall frequency power spectrum of the signal as well as some second order statistics of the temporal and spectral envelope of the sound but disrupted the characteristic higher spectro-temporal structure found in natural song.

Our study could not identify any selectivity for the BOS in any of the sub-regions of the primary auditory forebrain. However, it should be noted that in this study only the primary auditory regions were addressed. Hence, this does not exclude the existence of BOS selectivity emergence in other areas of the ascending auditory pathway. Furthermore, our fMRI results, supported by data from independently performed electrophysiological studies (Amin et al., 2007; Amin, Grace, & Theunissen, 2004; Lewicki & Arthur, 1996; Shaevitz & Theunissen, 2007), point to a clear hierarchical organization in terms of signal strength between L2a, L2b and CMM but show no obvious selectivity for temporal characteristics of the BOS in these regions.

After the implementation of BOLD fMRI in the zebra finch and the subsequent validation of the technique by traditional methods, these findings were
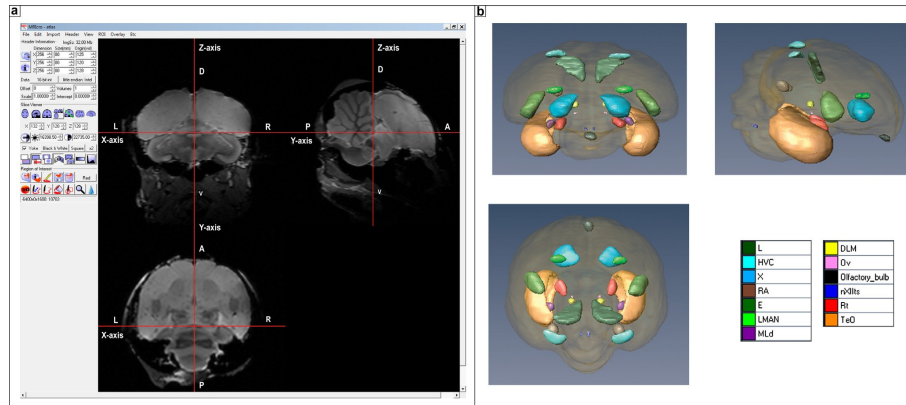
**Figure 2.5: Three-dimensional MRI atlas of the zebra finch brain.** (a) Atlas data set in the reference frame displayed in the MRIcro environment. Labels of axes were manually added for information but are not displayed by MRIcro software. The red cross is positioned on the 0 point corresponding to the mid-sagittal plane and the posterior commissure. A: anterior; D: dorsal; L: left; P: posterior; R: right; V: ventral. (b) Rendering of the whole brain and of delineated structures. Posterior view (top left), right view (top right), dorsal view (bottom left). L: Field L; HVC: used as a proper name; E: Nucleus Entopallialis; LMAN: Nucleus lateralis magnocellularis, pars lateralis; nXIIts: nXII Pars Tracheosyringealis; Rt: Nucleus Rotundus; TeO: Tectum Opticum. Redrawn from figures in Poirier et al. (2008).

the first contribution of BOLD fMRI to the understanding of auditory processing in the zebra finch brain. The functional difference between sub-regions L2a and L2b of the primary auditory center identified in this study were an important contribution of fMRI to the field because field L2 does not display increased IEG expression after exposure to song and electrophysiology only enables recording of a limited set of neurons and is thus not suited for such analysis.

## 2.4   Further technical refinements of the method

### 2.4.1   Improvements in acquisition protocol to allow imaging beyond the auditory system

Gradient Echo (GE) $T_2^*$-weighted as well as Spin Echo (SE) $T_2$-weighted contrasts can be used to perform BOLD fMRI. GE provides a higher contrast-to-noise ratio (CNR) and is thus more sensitive to variations in the BOLD response. Considering this, a GE-FLASH sequence was used for the first implementation of BOLD fMRI in the zebra finch brain. However, a major disadvantage of this sequence is that GE images contain large susceptibility artefacts which induce image distortions and signal dropout at high magnetic field strengths (e.g. 7 T which is used in our studies). Susceptibility artefacts are caused by the abundance of air cavities in the birds' cranial bones, thus

preventing the accurate measurement of the local BOLD response in some brain regions close to the skull. However, as our first zebra finch fMRI studies focused exclusively on primary and secondary auditory regions, located in the center of the brain, the signal loss associated with GE fMRI did not pose problems for these studies. Yet, since the further analysis of BOS selectivity required the analysis of BOLD response beyond the auditory regions involving the song control system (SCS), the need emerged to develop a technique that allowed for whole brain BOLD fMRI with less susceptibility artefacts. Therefore, SE BOLD fMRI was optimized for use in zebra finches in a new study at our lab (Poirier et al., 2010). Compared to GE fMRI, SE fMRI has the great advantage of providing signal from the entire brain as there is no signal dropout in the images. Another advantage of SE BOLD fMRI is its better spatial specificity (Harel, Lin, Moeller, Ugurbil, & Yacoub, 2006; Zhao, Wang, & Kim, 2004). Indeed, at high magnetic field, the intravascular component of the SE BOLD signal is reduced (due to a long echo time (TE)) and the extravascular component from large vessels is suppressed (by the 180deg refocusing pulse of the SE MRI sequence). The SE BOLD signal is thus dominated by an accurate extravascular signal originating from small vessels (Duong et al., 2003; Lee, Silva, Ugurbil, & Kim, 1999; Uludağ, Müller-Bierl, & Uğurbil, 2009). The main limitation of SE fMRI is its relatively weak sensitivity, requiring optimized sequences and optimized stimulation paradigms. Accordingly, a standard SE rapid acquisition relaxation-enhanced (RARE) protocol was used. Our study showed that, at 7 T, a TE of 60 ms provides a CNR and a signal-to-noise ratio sufficient to detect significant differences in BOLD responses triggered by different stimuli. Furthermore, a compromise had to be reached between imaging of the whole brain, a long repetition time (TR), an increased ramp time of the gradients to reduce background noise from the scanner and the maintenance of a sufficient temporal resolution for our stimulation paradigm. In order to achieve this goal, the ramp time was fixed at 600 $\mu$s allowing acquisition of 15 slices with a thickness of 0.75 mm$^3$ covering almost the whole brain within a TR of 2000 ms. This resulted in a temporal resolution of 8 s and an in plane resolution of 0.25 x 0.25 mm$^2$. With this sequence a purely $T_2$-weighted SE signal was obtained, which is characterized by a very good spatial specificity, a sensitivity sufficiently high to detect differential BOLD responses even beyond the primary and secondary auditory regions and a temporal resolution compatible with the stimulation paradigm (Poirier et al., 2011, 2009). For a detailed description of the SE-RARE fMRI protocol see Poirier et al. (2011, 2010); Van Ruijssevelt et al. (2012).

### 2.4.2 Progress in data analysis: voxel-based analysis and development of a zebra finch MRI brain atlas

In our first fMRI studies on songbirds, manual region of interest (ROI) analysis was applied for the statistical processing of the group data (Boumans, Gobes,

et al., 2008; Boumans et al., 2007; Boumans, Vignal, et al., 2008). However, as
our knowledge of the fMRI methods advanced and the acquisition parameters
became more optimized, the statistical analysis of the data was revised. Al-
though statistically powerful (signal intensity is averaged over all voxels in the
ROI, thus removing between-voxel variability), the ROI approach decreases
spatial resolution in the results and implicitly assumes that effects are identical
over all voxels in the ROI. This can be problematic if the ROI is large and het-
erogeneous (e.g. auditory region including the primary auditory region field
L, and the secondary auditory regions NCM and CM). Accordingly, other ap-
proaches which increase spatial resolution, were considered and voxel-based
analysis attracted our attention. At that time, voxel-based analysis was gain-
ing popularity in all fields. In the medical field it had mainly replaced manual
ROI analysis as it provided better tools to make statistical inferences at voxel
level. Shortly after, the next level of application, voxel-based meta-analysis,
was introduced. With this new approach, regional effects could be compared
across studies, providing quantitative summaries with good control and sta-
tistical error (Astrakas & Argyropoulou, 2010). To apply this method, a spa-
tial normalization or alignment of the acquired images to a common three-
dimensional coördinate space, is required. Digital, three-dimensional MRI
brain atlases were already available for humans (e.g. Lancaster et al., 2000;
Mazziotta et al., 2001), for macaques (e.g. Cannestra, Santori, Holmes, & Toga,
1997; Saleem & Logothetis, 2007), for baboons (Black, Snyder, Koller, Gado,
& Perlmutter, 2001), for rats (e.g. Schwarz et al., 2006; Schweinhardt, Frans-
son, Olson, Spenger, & Andersson, 2003) and for mice (e.g. Lin et al., 2003;
MacKenzie-Graham et al., 2004), but no such atlas existed for songbirds or any
other avian species. Hence, before implementation of voxel-based analysis in
the data processing of our zebra finch fMRI images, the first high-resolution
3D atlas in stereotaxic coordinates of a male zebra finch brain was developed
at our lab (Poirier et al., 2008) (Figure 2.5). The MRI atlas data set, the brain de-
lineation and the nuclei delineations were made freely available through the
Bio-Imaging Lab web site: http://tinyurl.com/ZFatlas and the atlas is already
proving its usefulness in different ongoing and published studies (e.g. Akuta-
gawa & Konishi, 2010; Maul et al., 2010; Poirier et al., 2009; Thode, Güttinger,
& Darlison, 2008).

    At present, the Bio-Imaging Lab also published a canary and a pigeon
brain MRI atlas (Güntürkün, Verhoye, De Groof, & Van der Linden, 2013;
Vellema, Verschueren, Van Meir, & Van der Linden, 2011), which can be eas-
ily adapted to match any surgical setup or histological protocol, and can thus
support many types of neurobiological studies, including anatomical, electro-
physiological, histological, explant, and tracer studies.

### 2.4.3 Effect of different anaesthetics on the BOLD response in the zebra finch's auditory system

To date, the majority of functional brain activation studies in animals using BOLD fMRI require that the animals be anaesthetized during the fMRI acquisition. Training animals to accustom them to the magnet environment and the scanner noise during the study periods is rather time-consuming and challenging and therefore rarely employed (Ferris et al., 2005, 2008; Sachdev et al., 2003). Moreover, anaesthesia minimizes stress-induced effects on the physiological responses of interest and facilitates animal handling (Van der Linden et al., 2007). The effect of anaesthesia both on the neural response and on the transfer function between the neural activity and the BOLD response measured in fMRI is an ongoing and important research topic. Anaesthetic level can profoundly change the BOLD response both because of changes in neural activity (Richards, 2002) and because of changes in the BOLD transfer function (Masamoto, Kim, Fukuda, Wang, & Kim, 2007). Various anaesthetics exist for use in zebra finches. Based on the findings about the compatibility of different anaesthetics with functional activation using haemodynamic fMRI in small animals, we chose to compare the effect of three anaesthetics acting on different neurotransmitter systems, on the BOLD response in a new study in zebra finches (Boumans et al., 2007). In this study, the effects of medetomidine, isourane and urethane were investigated. Medetomidine had already been used before with good results in our first functional imaging study in songbirds (Van Meir et al., 2005). Urethane had been used extensively in neurophysiological investigations of the avian auditory and vocal systems (reviewed in Theunissen et al., 2004) and its use allowed us to directly compare our results with those in the literature. Isourane is the most common anaesthetic in clinical applications as it has the great advantage of having relatively rapid recovery and minor side effects and thus has the highest potential for use in longitudinal studies. Consistent with our expectations and the findings of previous studies in small animals, the different anaesthetics induced different BOLD responses. One of the most prominent differences was the expanded area of activation when using isourane or urethane compared to medetomidine. Although not a lot has been published on the matter, urethane is considered toxic for laboratory animals including birds (e.g. Cowen, 1950) and is thus only suitable for terminal (acute) procedures and not for longitudinal studies. In addition, urethane is carcinogenic and poses a risk to personnel working with it (Schmidt, 2010). Consequently, isourane became our anaesthesia of choice for auditory BOLD fMRI studies in the zebra finch.

Throughout fMRI acquisition it is important to maintain a stable level of anaesthesia and thus dose. This dose is preferentially as low as possible in order to minimize the effects of the compound on the neural responses. Once an optimal dose is determined (for zebra finches the Bio-Imaging Lab uses mostly 1.5% isourane during fMRI acquisition) it is important that this dose is maintained for all animals in the study to maximize comparability of the acquired

data. The physiological status of anesthetized animals may also greatly inuence BOLD contrast. Monitoring of physiological parameters is therefore an important part in the design and performance of reliable fMRI experiments. Body temperature, for example, must be closely monitored and retained stable throughout acquisition using e.g. an MR compatible water bed heating system. Changes in body temperature may substantially affect the rate of chemical reactions and metabolic oxygen consumption and may also modulate the properties of the blood such as viscosity and affinity of hemoglobin for oxygen. These changes, in turn, have a great impact on the BOLD signal (Vanhoutte, Verhoye, & Van der Linden, 2006). In addition, it is important to retain stable global and cerebral hemodynamic parameters in the anesthetized animals. Levels of $CO_2$ and $O_2$ in the blood can have great inuence on CBF and CBV. As BOLD activation changes are in the range of only a few percent and based on changes in CBF and CBV, monitoring these blood gases to ensure stable levels is of great importance for accurate BOLD fMRI. Although not commonly used in small birds such as the zebra finch, it is possible to measure the partial pressure of oxygen ($pO_2$) non-invasively using a pulse oximeter clipped over e.g. the brachial artery (Nilson, Teramitsu, & White, 2005). Other more invasive methods can possibly also be used. However, the required blood sample volumes and MR compatibility limit their application to date.

## 2.5 BOLD response to subtle differences in song stimuli in the zebra finch song control and auditory system: new findings on BOS selectivity and the involvement of MLd

Selective responses to the BOS were initially observed using electrophysiological methods in the SCS, the network of forebrain regions involved in song production and learning (Doupe & Konishi, 1991; Janata & Margoliash, 1999; Margoliash, 1986; Margoliash & Fortune, 1992). Besides its appearance in the SCS, BOS selectivity was commonly expected to emerge at least at an intermediary level in the ascending auditory system (Prather & Mooney, 2004; Theunissen et al., 2004). Hence, numerous studies were designed to determine where this selectivity emerges. Because electrophysiology requires an a priori hypothesis about the location of the selectivity, most studies focused on the primary and secondary auditory regions. However, this hypothesis did not lead to any conclusive results (Bauer et al., 2008; Theunissen et al., 2004). Similarly, our previous fMRI study could not identify any selectivity for the BOS in any of the sub-regions of the primary auditory forebrain (Boumans et al., 2007). With the implementation of SE-BOLD fMRI, the production of the zebra finch MRI atlas and voxel-based statistics, we further investigated this
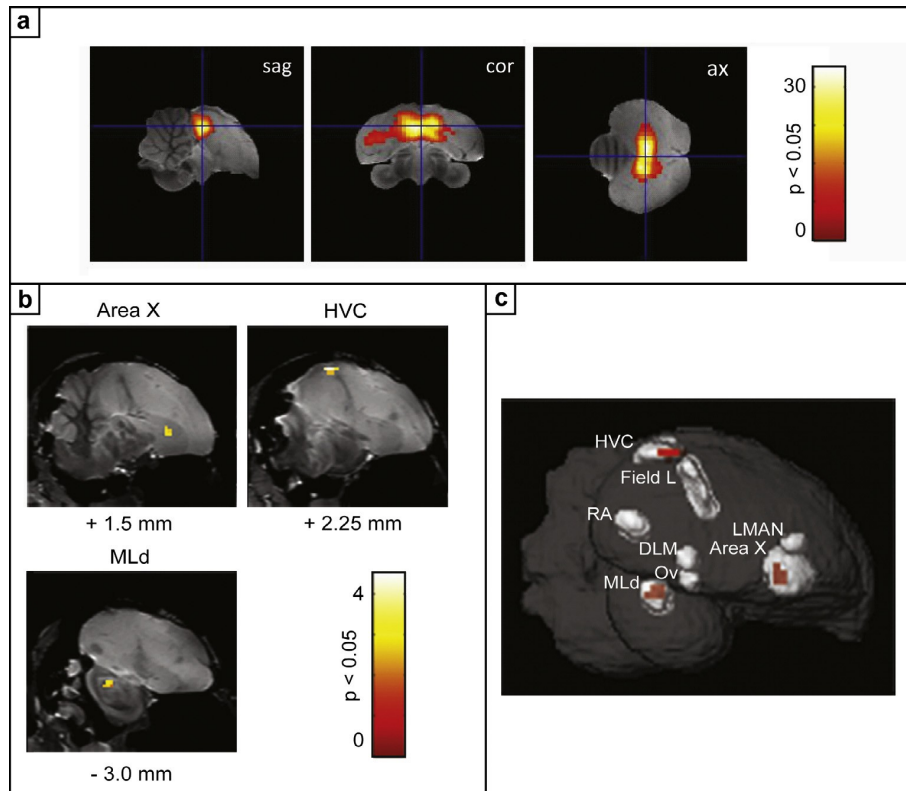
**Figure 2.6: Detection of neural selectivity for behaviorally relevant stimuli using fMRI.** Statistical parametric maps superimposed on high-resolution anatomical images from the zebra finch brain atlas (Poirier et al., 2008). T-values are colour coded according to the scale displayed in the figure and only voxels in which the t-test was found to be significant are displayed. (a) fMRI BOLD response in the primary auditory region, Field L, and adjacent secondary auditory regions evoked by stimulation with different types of conspecific song (BOS, familiar conspecific song and tutor song) compared to the rest condition (unpublished data). (b and c) Neural substrates of own-song recognition (BOS vs conspecific song) in the song control and auditory regions. Coordinates expressed in millimeters from the midline are indicated under each map. The + signs indicate that the visualized slices (and statistical results) are from the right hemisphere. (c) Superimposition of the statistical results (red clusters) to rendered images of the whole zebra finch brain with song control nuclei (HVC, RA, DLM, LMAN and Area X) and auditory nuclei (Field L, Ov and MLd). Redrawn from figures in Poirier et al. (2009).

question in the whole zebra finch brain (Poirier et al., 2009). Consistent with the neural electrophysiological literature, our study, which compared neural activation elicited by the BOS and a familiar conspecific song, revealed BOS selectivity in large neuronal populations of HVC (and the HVC shelf) and part of area X (Figure 2.6). Additionally, this study revealed the presence of activity-based selectivity to the BOS within the ascending auditory pathway at the level of the midbrain (nucleus mesencephalicus lateralis pars dorsalis (MLd); homolog to the inferior colliculus) where it had never been investigated before. Surprisingly, this selectivity was lateralized towards the right side, a finding reminiscent of the neural right lateralization of self-voice and self-face recognition in humans (Keenan, Nelson, O'Connor, & Pascual-Leone, 2001; Nakamura et al., 2001). A previous fMRI study already suggested that activation by songs could be lateralized because some effects were observed on the right side of the brain only, although this study did not report any direct comparison of selectivity between the two hemispheres (Voss et al., 2007). Interestingly, selectivity for conspecific over heterospecific songs was also observed in the MLd but in this case it was lateralized to the left side of the brain (Poirier et al., 2009). Therefore, it is important that future electrophysiological and IEG experiments systematically measure and report results from each hemisphere separately.

## 2.6 New findings on the role of MLd in tutor song selectivity and memory

The use of the zebra finch MRI atlas together with an increasing optimization of the methods for spatial registration and statistical processing of zebra finch fMRI data enabled the detection of increasingly subtle neural activations in response to acoustically similar song stimuli. This sensitivity was needed in order to discern tutor song selectivity from selectivity for the BOS and familiar conspecific songs without introducing caveats due to acoustic differences. In a study addressing the neural substrates of tutor song memory (reported in Chapter 3 of this dissertation) songs with different behavioral relevance were presented to zebra finches, while restricting acoustic differences between stimuli. With this experimental design, the fine line between BOS and tutor song responses in the adult zebra finch brain was explored in a single fMRI experiment.

In the study described in Chapter 3 we presented 17 adult male zebra finches with their own song, their tutor's song and a highly familiar and acoustically close conspecific song while collecting whole-brain fMRI images. We thus demonstrated that in adult zebra finches the right MLd selectively responds to the copied song of the tutor and the BOS. The selective responses to tutor song and BOS were shown to be distinct. Furthermore, the amplitude of the selective signal was shown to be positively correlated with the strength

of vocal learning, measured by the amount of song that experimental birds copied from the adult model.

Although a form of tutor song memory has been hypothesized in the telencephalic auditory region NCM by several authors (Bolhuis & Gahr, 2006; Gobes & Bolhuis, 2007; London & Clayton, 2008; Phan, Pytte, & Vicario, 2006), our data did not show any tutor song selective BOLD activity in this region. These different findings might result from the use of anaesthesia during fMRI acquisition or from methodological differences. Whereas IEG expression in response to song represents the overall neural response to a stimulus, differences between stimuli (e.g. differential activation evoked by the tutor song compared to a conspecific song) are assessed when processing fMRI data. The presence of selective responses which cannot be explained by acoustic differences between the stimuli, strongly suggests that MLd is part of the neural substrates of tutor song memory, because the memorization and copying of this song early in life is the one attribute discerning the tutor song from the other stimuli. Reinforcing this interpretation, the strength of tutor song selectivity was found to be positively correlated with the amount of song that the experimental birds copied from their tutor. This correlation suggests that birds that formed a strong memory of their tutor's song later produced an accurate copy of this song.

With this study we were able to confirm the presence of BOS selectivity in MLd (Poirier et al., 2009) and to show for the first time neural selectivity for the tutor song in the auditory midbrain. The finding that early sensory experience can generate long-lasting memories in a midbrain structure provides additional evidence to the emerging view that experience-dependent plasticity is not limited to cortical structures (Tzounopoulos & Kraus, 2009; Xiong, Zhang, & Yan, 2009). For a detailed description of this study, please refer to Chapter 3.

## 2.7   fMRI of zebra finch during ontogeny

Previous paragraphs have shown the ability of fMRI to demonstrate activation of songbird brain regions in response to complex song stimuli exhibiting minor acoustic differences. With the increasing body of research on neural song selectivity in adult birds, both from fMRI studies and studies employing other methods, it has now become possible to study song processing in juvenile birds while they are still in the process of learning their song. Studying juvenile birds with a non-invasive method like fMRI creates interesting possibilities with regard to the study of song learning and its comparison to human language learning studies.

fMRI studies to juvenile birds using a longitudinal in vivo approach can add to the current knowledge on the neural underpinnings of zebra finch song learning by following the development of the neural mechanisms for song processing and production that have been previously found in adult

birds. Because birds can be followed throughout their lives, any changes in the neural correlates for song perception that might occur between juvenility and adulthood can be studied within the same group of birds. These differences between juvenile and adult neural activations within the same birds can tell us something about the functional brain changes that might underlie the sensitive period for song learning and song crystallization. Moreover, whole-brain fMRI in juvenile zebra finches can potentially enable mapping of the neuro-functional correlates of all stages of song learning and the development of the entire brain within-subject.

Results of the first fMRI study on juvenile zebra finches, conducted in the Bio-Imaging Lab, are discussed in Chapter 4 of the present thesis. In order not to disturb normal rearing conditions, we chose to start test imaging the juvenile zebra finches from 20 days post hatching (DPH) or after edging if this did not happen by day 20. At this age, the reaction of the juvenile birds to isourane anaesthesia was comparable to that of an adult bird. The juveniles showed a stable respiration pattern at the same low dose as adult birds received in earlier experiments ($\pm 1.5\%$). Although it is possible to measure juvenile zebra finches from the age of 20 DPH, the zebra finch brain reaches its adult size around 30 DPH. Therefore, images can be more easily registered to the zebra finch MRI atlas after the age of 30 DPH which makes it possible to compare data of the same birds at different ages. Taking this factor into account, Chapter 4 reports within-subject results on juvenile zebra finches measured at different time points from 30-120 DPH while listening to tutor song, conspecific song and heterospecific song.

## 2.8 New findings on BOS selectivity: involvement of the noradrenergic system

Previous studies showed thus that the songbird brain is able to discriminate between the BOS and other conspecific songs (e.g. Poirier et al., 2009). Since song is a learned behavior, the development of BOS selectivity necessarily involves experience-dependent mechanisms. Accordingly, brain regions sensitive to self-generated vocalizations are likely to mediate the auditory feedback critical for song learning and song maintenance. The presence of high concen- trations of norepinephrine (NE) and NE receptors in the song control and auditory nuclei, associated with the observation that the development of the noradrenergic innervation closely parallels song learning, have led to the hypothesis that NE might be involved in the control of song production, perception and learning (for a review, see Castelino & Schmidt, 2010). Additionally, electrophysiological evidence indicates that NE can suppress BOS responsiveness in song control regions (Cardin & Schmidt, 2004; Dave et al., 1998). Hence, a new study at our lab was set up to further investigate the role of NE on BOS selectivity. To manipulate the NE level in the zebra finch brain in

order to investigate the functions of the central noradrenergic system in zebra finches, we systematically injected the birds with the selective noradrenergic neurotoxin DSP-4, a neurotoxin that induces degeneration of the two main noradrenergic cell groups innervating the telencephalon, the locus coeruleus and nucleus subcoeruleus ventralis (Castelino & Ball, 2005; Waterman & Harding, 2008).

The administration of DSP-4 and thus the suppression of noradrenergic activity was shown to unmask BOS selectivity and induce a BOS selective BOLD signal in the dorsal NCM. This finding is in accordance with the electrophysiological results mentioned above and suggests that NE might play a role in the suppression of BOS selectivity. It is also known that NE levels in the telencephalon are low in sleeping birds compared to alert birds. This might explain the high levels of BOS selectivity found in zebra finches during sleep (for a review, see Margoliash & Schmidt, 2010).

## 2.9 Developments and applications of the technique in other research groups

The studies reviewed above summarize the successful implementation of BOLD fMRI in the zebra finch brain as developed in the Bio-Imaging Lab as well as its validation using different methods including optical imaging and IEG expression. They also demonstrate the applicability of this technique to the detection of neural selectivity for behaviorally relevant stimuli and thus illustrate the importance of BOLD fMRI for the study of the neural substrates of zebra finch song perception and song learning. Although the songbird brain has been studied extensively using other methods, BOLD fMRI has only recently been applied to the songbird model in other research groups than ours. With their first study published in 2007, Voss and colleagues have employed GE BOLD fMRI in the songbird model using a 3 T human MRI system in combination with a single transmit/receive coil (Voss et al., 2007).

Using GE imaging, relatively high signal strength can be obtained, but the specificity is reduced due to a lower spatial resolution compared to imaging in a 7 T small animal system. Despite the sensitivity to susceptibility artefacts of the GE sequences used in their studies, Voss and colleagues obtained interesting results by using a lower field strength and a reversed gradient distortion correction method in combination with a focus on the Field L/NCM complex, which is relatively unaffected by these artefacts. In their first zebra finch fMRI study, Voss and colleagues exposed male zebra finches to tones and song with varying behavioral relevance (Voss et al., 2007). Analysis of the area of activation in the caudal telencephalon in response to the different stimuli showed a wider extent of BOLD activation for BOS and TUT compared to conspecific song and tones. Furthermore, the right hemisphere showed an overall better discrimination between stimuli and a difference between the activation pat-

terns for BOS and TUT was only found in the right hemisphere. These results are in line with our finding that only the right auditory midbrain shows BOS as well as tutor song selectivity and might indicate that the right hemisphere plays an important role in tutor song memory and auditory feedback learning.

In the same research group, Maul et al. (2010) demonstrated that isolated females and both males and females with song exposure show more focal BOLD activity in response to conspecific song with the caudal telencephalon compared to isolated males. These results suggest that, in contrast to males, female zebra finches do not need song exposure in order to develop song-specific neural responses (Maul et al., 2010). In a recent study Voss and colleagues employed voxel-based statistics in order to directly assess differences in auditory forebrain activations between normal birds and birds that show stuttered syllables in their song (Voss et al., 2010). They thus showed a reduced tutor song related activation in Field L in birds with stuttered syllables in their song compared to birds with a normal song. This finding indicates that the birds with stuttered syllables have a deficient memory of their tutor's song and thus adds to the idea that the ascending auditory pathway plays a more important role in auditory memory formation and song learning than previously thought.

All the experiments of Voss and colleagues were conducted using diazepam sedation rather than isourane anaesthesia, used in most of our studies. However, this difference is not expected to have large effects on the results, because electrophysiological responses of neurons have been shown to be similar under anaesthesia and sedation in songbirds (Cardin & Schmidt, 2003). Moreover, song selectivity in the auditory midbrain is unlikely to be inuenced by anaesthetic state, because anaesthesia has been shown not to modulate spectral tuning or population representations of song in auditory midbrain neurons (Schumacher, Schneider, & Woolley, 2011).

## 2.10   Conclusions and future prospects

The results discussed above confirm the specialization of the caudal telencephalon for the processing of conspecific songs and the existence of neural selectivity for the BOS and tutor song in the songbird brain. BOLD fMRI has been shown to be a valuable technique for the mapping of neural mechanisms of song learning because it is a non-invasive technique that can be used in longitudinal studies. In addition, it allows for a whole-brain approach. The growing body of zebra finch fMRI research also demonstrates its applicability to cognitive research questions. A future effort which should be considered in order to further extend fMRI studies beyond auditory processing, a neural process which remains active under anaesthesia (Theunissen et al., 2004), is the use of awake subjects habituated to the imaging protocol. This might offer new research possibilities where birds can actively participate in fMRI experiments and where the attention of the subjects can be experimentally

controlled.

At present, due to its non-invasiveness, BOLD fMRI can be increasingly exploited as a complementary method to inform local and invasive methods. Recent advancements in unraveling the genetic background of singing behavior have led to a better understanding of the genes that are linked to vocal learning in birds (Haesler et al., 2007; Hilliard, Miller, Fraley, Horvath, & White, 2012). This raised questions about how these genes affect neurodevelopment in relation to song learning. Developmental studies employing BOLD fMRI in transfected or, in the future, transgenic zebra finches, might uncover the neural mechanisms linking the genetic background to behavior. Furthermore, BOLD fMRI is a method very well suited for the advanced mapping of gene transfection targets. Because MRI can also be performed in small rodents, this technique can provide understanding of the similarities and differences in the neural correlates of birdsong and ultrasonic vocalizations in mice, which have both been associated with the FoxP2 gene, a gene which is hypothesized to be involved in vocal communication in humans (e.g. Fischer & Hammerschmidt, 2011; Haesler et al., 2007). However, at present auditory fMRI has not yet been performed in mice. Currently, the only studies reporting on the mapping of sound-evoked activity in the mouse brain are based on the use of manganese-enhanced MRI (e.g. Yu et al., 2008).

In contrast to the invasive methods classically used in songbird studies, BOLD fMRI is a method widely employed in the field of human speech and language research. For this reason, a meaningful comparison can be made between human speech perception and birdsong perception fMRI research to address questions about the origin of human vocal learning. Although most fMRI research on human adults is carried out in awake subjects, there is an increasing body of research using the haemodynamic method NIRS to address speech perception in human infants during sleep.Furthermore, the development of MRI imaging during wakefulness could add to the applicability of bird fMRI to human speech perception research.

To conclude, BOLD fMRI in zebra finches has been widely validated and has advanced beyond the level of simple discrimination. Stimuli with increasingly fine-grained manipulations and acoustic differences can be used to study complex neural mechanisms and their development in the entire songbird brain. By combining songbird BOLD fMRI with a wide range of other techniques, an integrative research framework can be created to exhaustively map the neural correlates of vocal learning in songbirds as well as other species.

## 2.11 Acknowledgements

CHAPTER 3

---

# Representation of early sensory experience in the adult auditory midbrain: implications for vocal learning

---

*Vocal learning in songbirds and humans occurs by imitation of adult vocalizations. In both groups, vocal learning includes a perceptual phase during which juveniles birds and infants memorize adult vocalizations. Despite intensive research, the neural mechanisms supporting this auditory memory are still poorly understood. The present functional MRI study demonstrates that in adult zebra finches, the right auditory midbrain nucleus responds selectively to the copied vocalizations. The selective signal is distinct from selectivity for the bird's own song and does not simply reflect acoustic differences between the stimuli. Furthermore, the amplitude of the selective signal is positively correlated with the strength of vocal learning, measured by the amount of song that experimental birds copied from the adult model. These results indicate that early sensory experience can generate a long-lasting memory trace in the auditory midbrain of songbirds that may support song learning.*

## 3.1 Introduction

Songbirds share with humans the ability to learn their vocalizations (Bolhuis, Okanoya, & Scharff, 2010; Doupe & Kuhl, 1999; Wilbrecht & Nottebohm, 2003). Like human babies need to be exposed to adult speech to develop a

normal vocal repertoire, juvenile songbirds need to be exposed to adult conspecific vocalizations to develop a normal song (sensory phase). Then, during a subsequent sensori-motor phase, they use auditory feedback to progressively match their own developing vocalizations to the memorized adult model (called tutor song) (Konishi, 1965). Learning by imitation requires first to compare the motor performance with the object of imitation and then to correct for potential errors. It has long been hypothesized that the anterior forebrain pathway of songbirds, a circuit driving vocal variability in juveniles and adults (Kao, Doupe, & Brainard, 2005; Ölveczky, Andalman, & Fee, 2005; Scharff & Nottebohm, 1991), participates in both vocal error detection and error correction (Brainard, 2004).

While the role of the anterior forebrain pathway in generating a corrective premotor bias has been recently confirmed (Andalman & Fee, 2009), a growing number of studies point to the ascending auditory pathway to be the main neural substrate of tutor song memory (Bolhuis & Gahr, 2006; Gobes & Bolhuis, 2007; London & Clayton, 2008; Phan et al., 2006) and feedback-dependent error detection (Keller & Hahnloser, 2009; Lei & Mooney, 2010). However, if the auditory system supports the comparison between the bird's own song and a memory trace of the tutor song in order to detect vocal errors, one would expect to find bird's own song and tutor song selective signals in some of the auditory nuclei (Margoliash & Schmidt, 2010). While significant bird's own song selective responses have been recently found in the auditory midbrain (Poirier et al., 2009) and the auditory thalamus (Lei & Mooney, 2010), evidence for tutor song selective responses in the ascending auditory pathway is still missing. The goal of this study was thus to look for tutor song selectivity in the auditory system, using blood oxygen level-dependent (BOLD) functional MRI (fMRI), a technique commonly used on humans and recently adapted to songbirds (Van der Linden, Van Meir, Boumans, Poirier, & Balthazart, 2009). Such selectivity was found in the right auditory midbrain.

## 3.2    Materials and Methods

### 3.2.1    Ethical Statement

All experimental procedures were performed in accordance with the Belgian laws on the protection and welfare of animals and were approved by the ethical committee of the University of Antwerp, Belgium (EC no. 2009/21). All fMRI recordings were performed under isoflurane anesthesia and all efforts were made to minimize suffering and anxiety.

### 3.2.2    Subjects

Twenty adult male (mean age 24 months, range 10-41 months) zebra finches (Taeniopygia guttata) recruited from the breeding colony of the Max Planck

Institute for Ornithology (Seewiesen, Germany) were used in this experiment. Birds were raised by their parents from 0 to 7 days post hatching (DPH), by their mother from 8 to 34 DPH and were kept alone from 35 to 42 DPH. The birds were then housed singly with one adult male tutor from 43 to 100 DPH (one-to-one paradigm). Thirteen different tutors were used in the present experiment. These tutors previously learned their own song from one of three song models via tape playback. Song data collected on the experimental birds and their tutors indicate that the three song models elicited similar amount of song copy. Following tutoring (after 100 DPH), the experimental birds were housed together, first in aviaries then in large cages. Birds were maintained throughout the experiment under a 12h light: 12h dark photoperiod and had access to food, water and baths ad libitum.

### 3.2.3 Song Recording and Analysis

Prior to the fMRI experiment, each experimental bird was placed alone during 48 hours in a soundproof chamber and its song was recorded using the Sound Analysis Pro (SAP) 2.0 software (Tchernichovski, Nottebohm, Ho, Pesaran, and Mitra (2000); http://soundanalysispro.com/). Acoustic similarity between songs was assessed using the similarity score implemented in SAP. This measure is based on five acoustic features: pitch, frequency modulation, amplitude modulation, goodness of pitch and Wiener entropy and comprises two components: the percentage of similarity, measuring at a large scale (70 ms) the amount of sound shared between two songs and the 'accuracy', measuring the local, fine grained (10 ms) similarity (for more details, see SAP user manual, available at http://soundanalysispro.com/). The final score corresponds to the product of these two components. The computation of this similarity score was done by selecting one song as a reference (asymmetric measurement).

To measure the vocal learning strength of each experimental bird, we selected the tutor song as the reference song, and compared the song of the tutee to this reference. This procedure was repeated 100 times, comparing 10 different exemplars of the tutor song with 10 different exemplars of the tutee song; the mean value was used. For measuring the acoustic similarity between stimuli used in the fMRI experiment (see below), there was no reason to choose one stimulus as a reference rather than the other one. For each pair of stimulus, we thus computed the similarity score twice, first using one stimulus of the pair as the reference, then using the other stimulus as the reference and finally computed the mean between the two indices.

### 3.2.4 fMRI stimuli

For each experimental bird, three familiar songs were used as stimuli in the fMRI experiment: the bird's own song (BOS), the tutor song (TUT) and a conspecific song (CON). The conspecific song came from an adult bird housed

during several weeks in the same aviary or cage as the experimental bird after the end of the learning phase, (i.e. after 100 DPH). This adult bird had been previously raised by a tutor, which had learnt to copy the same song model than the tutor of the experimental bird (Figure 3.1). As a result, the CON stimulus was thus not only familiar to the experimental bird but also acoustically close to its own song and its tutor song. For each bird, stimuli corresponded to one song exemplar of each category (BOS, TUT and CON), picked up randomly from the 10 exemplars used for computing the learning strength value (see above). Measures of acoustic similarity revealed no significant difference between the three stimuli (Repeated measure one-way ANOVA: $F = 0.98$, $p = 0.39$). Post-hoc paired t-tests confirmed the absence of significant difference between each pair of stimulus (TUT/CON similarity vs. TUT/BOS similarity: $t = 0.48$, $p = 0.64$; TUT/CON similarity vs. BOS/CON similarity: $t = 1.3$, $p = 0.21$; TUT/BOS similarity vs. BOS/CON similarity: $t = 1.1$, $p = 0.28$).

### 3.2.5   Experimental setup and design

During the experiment, birds were continuously anaesthetized with $1.5\%$ isoflurane. Auditory stimuli were played back at a mean intensity (in terms of Root Mean Square) of 70 $dB$ through small loudspeakers (Visation, Germany) from which magnets were removed. An equalizer function was applied to the stimuli using WaveLab software (Steinberg, Germany) to correct for enhancement of frequencies between 2500 and 5000 $Hz$ in the magnet bore (see Poirier et al., 2010). Stimulus delivery was controlled by Presentation 0.76 software (Neurobehavioral Systems Inc., Albany, CA, USA).

During fMRI acquisition, the three stimuli were randomly presented in an ON/OFF blocked design where 16 s. stimulation (ON blocks) and 16 s. rest periods (OFF blocks) were alternated. Each ON block included repetitions of the same stimulus interleaved with silent periods. The duration of the silent periods was adjusted in each bird to match the amount of song and silence between stimuli (mean song duration: 11.2 s for each stimulus; mean silence duration: 4.8 s). The experiment consisted in 93 ON blocks (31 per stimulus) and 93 OFF blocks. During each block, 2 magnetic resonance images were acquired, resulting in 62 images per stimulus and per subject.

### 3.2.6   fMRI acquisition

BOLD fMRI images were acquired using a 7T Pharmascan system (Bruker, Erlangen, Germany). Details about this system and the coils used for the experiment can be found in Boumans et al. (2007). For each bird, a time series of 372 T2-weighted rapid acquisition relaxation-enhanced (RARE) Spin Echo (SE) images (Echo time (TE)/Repetition time (TR): 60/2000 ms; RARE factor: 8; Field of View: 16 x 16 mm) was acquired. Images comprised 15 slices (in-plane resolution: 250 x 250 $\mu m^2$) with a slice thickness of 750 $\mu m$ and an inter-
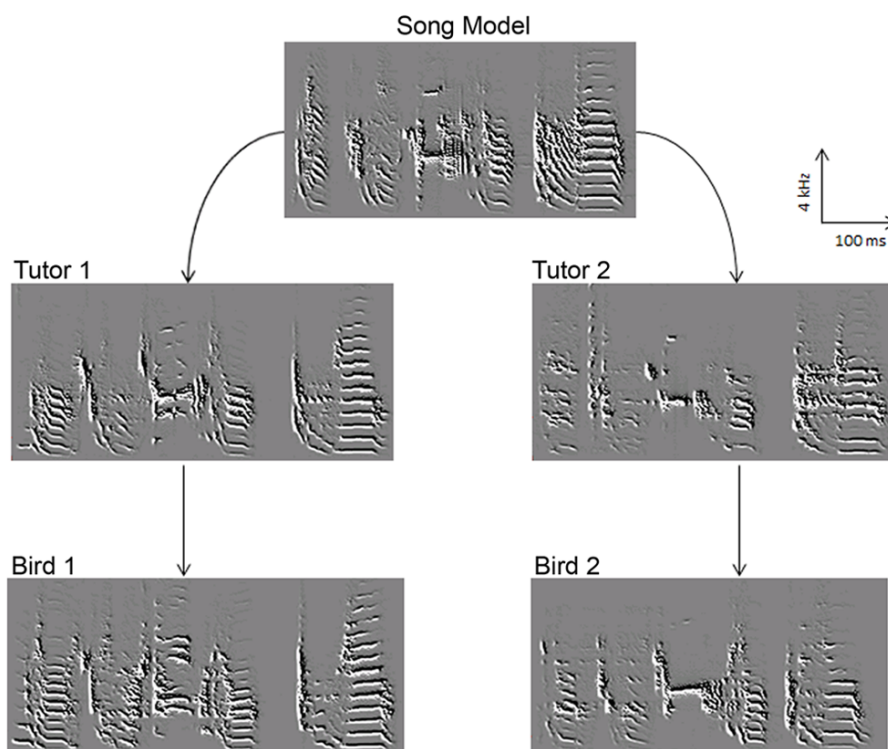
**Figure 3.1: Sonograms illustrating the song tutoring protocol for two experimental birds (Bird 1 and 2).** Tutors 1 and 2 learned their song from the same song model (via tape playback) while experimental birds 1 and 2 learned their song by being housed with respectively tutor 1 and tutor 2 (one-to-one paradigm). As a result, songs of Bird 1 and 2 were acoustically close. During the fMRI experiment, bird 1 was exposed to the song of bird 1 (BOS), the song of Tutor 1 (TUT) and the song of Bird 2 (CON).

slice gap of 50 $\mu m$, covering the whole brain. Following the fMRI acquisition, a high-resolution anatomical three-dimensional (3D) SE RARE image (voxel size 125 $\mu m^3$; TE/TR: 60/2000 ms; RARE factor: 8; Field of View: 16 x 16 mm) was acquired for each bird.

### 3.2.7   Image processing

Data processing was carried out using SPM8 (Wellcome Trust Centre for Neuroimaging, London, UK; http://www.fil.ion.ucl.ac.uk/spm/). To enable an accurate localization of the functional activations, the high-resolution anatomical 3D images of each subject were normalized to the MRI atlas of the zebra finch brain (Poirier et al., 2008). Each fMRI time series was realigned to correct for head movements, co-registered to the high-resolution 3D image of the same bird and up-sampled to obtain a resolution of 125 x 125 x 400 $\mu m$, as classically done in fMRI data processing. These steps resulted in a good correspondence between the fMRI data and the anatomical data from the atlas. Finally, the fMRI images were smoothed with a Gaussian kernel (width of 500 x 500 x 800 $\mu m$).

### 3.2.8   Statistical analysis

Statistical analysis of the fMRI data was performed at the subject and group level in SPM8, using the General Linear Model. Data were modeled as a boxcar and filtered with a high-pass filter of 352 seconds. Model parameters were then estimated using a classical restricted maximum likelihood algorithm. Subject-level analyses were performed to identify the mean effect [All stimuli minus rest] in each individual subject. These analyses revealed a bilateral positive BOLD signal in the auditory telencephalic regions (Figure 3.2) of 17 birds over 20, a success rate similar to the one obtained in our previous spin-echo fMRI experiments (Poirier et al., 2011, 2009, 2010). A bilateral response to the stimulation paradigm in the auditory regions confirms that the stimulation has been processed by the auditory system and was therefore used as an inclusion criterion. The subsequent analyses were thus only performed on these 17 birds, data from the 3 remaining birds being discarded.

The effect of [each stimulus minus rest] of each subject was then entered in a group-level random effect analysis. The mean effect [All stimuli minus rest] at the group level revealed a positive BOLD response not only in the auditory telencephalic regions but also in the dorsal part of the lateral mesencephalic nucleus (MLd), the main auditory midbrain nucleus. In order to increase the sensitivity of the statistical analyses, we focused on two pre-defined regions of interest in each hemisphere: MLd, where bird's own song selectivity has been previously found (Poirier et al., 2009) and the caudomedial nidopallium (NCM) (Figure 3.3), a telencephalic auditory region previously shown to be involved in tutor song memory (Bolhuis & Gahr, 2006; Gobes & Bolhuis, 2007; Phan et al., 2006). MLd could be clearly identified and delineated on the zebra
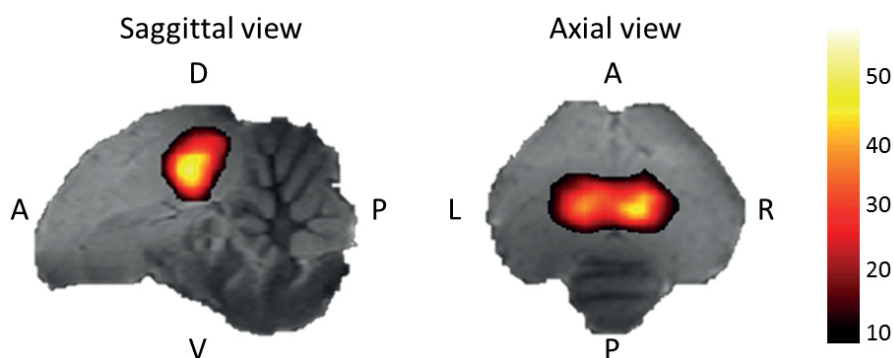
**Figure 3.2: Statistical maps of BOLD activation induced by all stimuli together.** Results (compared to Rest) are superimposed on anatomical sagittal and axial images coming from the MRI zebra finch atlas. T-values are color-coded according to the scale displayed on the right side of the figure. Only significant voxels (one-tailed t-test, $p < 0.05$, corrected at the whole brain level) are displayed. L: left, R: right, D: dorsal, V: ventral, A: anterior, P: posterior.

finch atlas (Poirier et al., 2008). NCM was delineated using Field L as anterior border, the cerebellum as posterior border and the lateral ventricle as ventral and dorsal borders. The lateral boundaries of NCM are not defined from a cyto-architectural point of view. In accordance with previous functional studies (Bolhuis, Hetebrij, Den Boer-Visser, De Groot, & Zijlstra, 2001; Bolhuis, Zijlstra, Den Boer-Visser, & Van der Zee, 2000; Chew, Mello, Nottebohm, Jarvis, & Vicario, 1995; Stripling, Volman, & Clayton, 1997; Terpstra, Bolhuis, & Den Boer-Visser, 2004), we included the three 0.4 mm-thick slices covering brain tissues between 0.2 mm and 1.4 mm from the midline in each hemisphere.

Statistical differences between stimulus-evoked BOLD signals were assessed in each voxel of the predefined regions using a one-way repeated measure ANOVA (F-tests) followed by post-hoc one-tailed paired t-tests. P-values were corrected for multiple tests using the Family Wise Error method based on the Random Field Theory (Worsley et al., 1996). In addition, an extent threshold was applied to the results: activations had to consist of a cluster of at least 5 significant contiguous voxels (corrected $p$-value $< 0.05$) to be considered statistically significant. Reflecting the voxel basis of the analysis, results are reported by the highest voxel $F/t$-value within each cluster (Fmax/tmax) and the associated voxel $p$-value. Regression analyses were also performed to assess potential correlations between the amplitude of differential fMRI signals ([BOS minus CON] and [TUT minus CON]) and various behavioral measures. In MLd, these analyses were performed by taking the mean fMRI signal averaged over the contiguous voxels in which a significant differential fMRI signal was first demonstrated. When applied to a brain region which can be reasonably assumed to be homogeneous, this procedure is more representative of data than a voxel-based analysis (i.e. correlation analysis performed in each individual significant voxel). Note however that a voxel-based analysis has
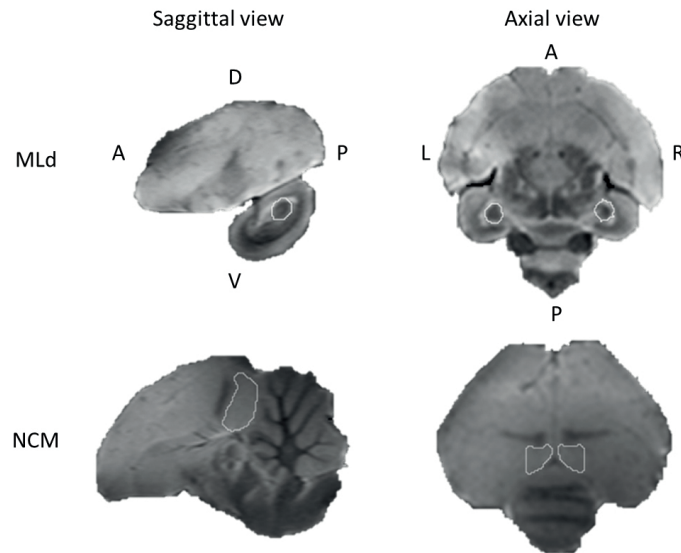
**Figure 3.3: Illustration of the predefined regions of interest on sagittal and axial anatomical images.** The anatomical images come from the zebra finch MRI atlas. L: left, R: right, D: dorsal, V: ventral, A: anterior, P: posterior.

also been performed and provided similar results (not described in the present manuscript).

In NCM, because the main effect of the ANOVA did not yield significant results, a correlation analysis between non-significant differential fMRI signals and learning strength was not meaningful. However, because previous authors reported a correlation between TUT-induced immediate early gene expression and learning strength in NCM (Bolhuis et al., 2001, 2000; Terpstra et al., 2004), we tested for potential correlation between [TUT minus Rest] and learning strength. Here, because the comparison [TUT minus Rest] was found significant in most part of NCM, we used a voxel-based approach. This approach was considered more relevant than using the mean fMRI signal averaged over all the NCM contiguous significant voxels because of the big size of NCM and the numerous studies suggesting that NCM comprises anatomically and functionally different sub-regions (e.g. Pinaud, Fortes, Lovell, & Mello, 2006; Pytte, Parent, Wildstein, Varghese, & Oberlander, 2010; Terpstra et al., 2004). Subsequent correlation analyses between learning strength and respectively [BOS minus Rest] and [CON minus Rest] were then limited to the small part of NCM where a correlation between [TUT minus Rest] and learning strength had been found, and were performed on the mean fMRI signal averaged over the contiguous voxels of this small region.

## 3.3 Results

### 3.3.1 Behavioral results of song tutoring

On average, the one-to-one tutoring protocol induced significant learning of the tutor song from the tutees: the mean learning strength, measured by the SAP similarity score including large-scale and fine-grained similarity, was of 48% (SE = 3.2), whereas the similarity of the tutee song with songs of other experimental birds heard only after what is supposed to be the end of the learning period (100 DPH) was of 28% (SE = 1.5). When learning strength was assessed by the SAP similarity score restricted to large-scale similarity, the mean value was 67%, which is within the range of what is accepted as normal tutor song copy; for instance, birds trained with tape recordings of adult songs were previously reported to have a large-scale SAP similarity score of 61% while birds raised with their parents had a score of 71% (Phan et al., 2006).

### 3.3.2 Brain responses in MLd

Right and left MLd were significantly positively activated by the three song stimuli BOS, TUT, and CON (Figure 3.4; Left MLd: [BOS minus Rest]: tmax = 6.7, $p < 0.0001$; [TUT minus Rest]: tmax = 4.5, $p = 0.001$; [CON minus Rest]: tmax = 5.2, $p = 0.0001$; Right MLd: [BOS minus Rest]: tmax = 6.9, $p < 0.0001$; [TUT minus Rest]: tmax = 6.7, $p < 0.0001$; [CON minus Rest]: tmax = 6.0, $p < 0.0001$). Significant differences in term of BOLD response amplitude elicited by different stimuli were found in right MLd (Fmax = 10.3, $p = 0.01$) but not in left MLd (Fmax = 3.2, $p = 0.35$). Post-hoc paired t-tests in right MLd revealed that the main effect was due to a greater activation induced by BOS and TUT compared to CON ([TUT minus CON]: tmax = 4.1, $p = 0.005$; [BOS minus CON]: tmax = 4.0, $p = 0.005$; [TUT minus BOS]: tmax = 1.1, $p = 0.57$).

Besides the fact that the mean acoustic similarity was not significantly different between each pair of stimuli (see Materials and Methods), we further examined whether the amplitude of the differential activations was correlated with the acoustic similarity between the stimuli. None of the correlations was significant (Figure 3.5; [TUT minus CON] vs. TUT/CON similarity: $R^2 = 0.14$, $p = 0.15$; [BOS minus CON] vs. BOS/CON similarity: $R^2 = 0.04$, $p = 0.44$; [TUT minus BOS] vs. TUT/BOS similarity: $R^2 = 0.03$, $p = 0.51$), excluding the acoustic similarity between the stimuli as the mere explanation for the amplitude of the differential activations.

We then looked whether the amplitude of the TUT and BOS selective signals (defined respectively as [TUT minus CON] and [BOS minus CON] BOLD responses) could reflect the amount of sound each experimental bird copied from its tutor (learning strength). This analysis revealed a significant positive correlation between TUT selectivity and learning strength (Figure 3.6; $R^2 = 0.36$, $p = 0.01$) as well as between BOS selectivity and learning strength ($R^2 = 0.25$, $p = 0.04$). Finally, we tested for potential correlations between the am-
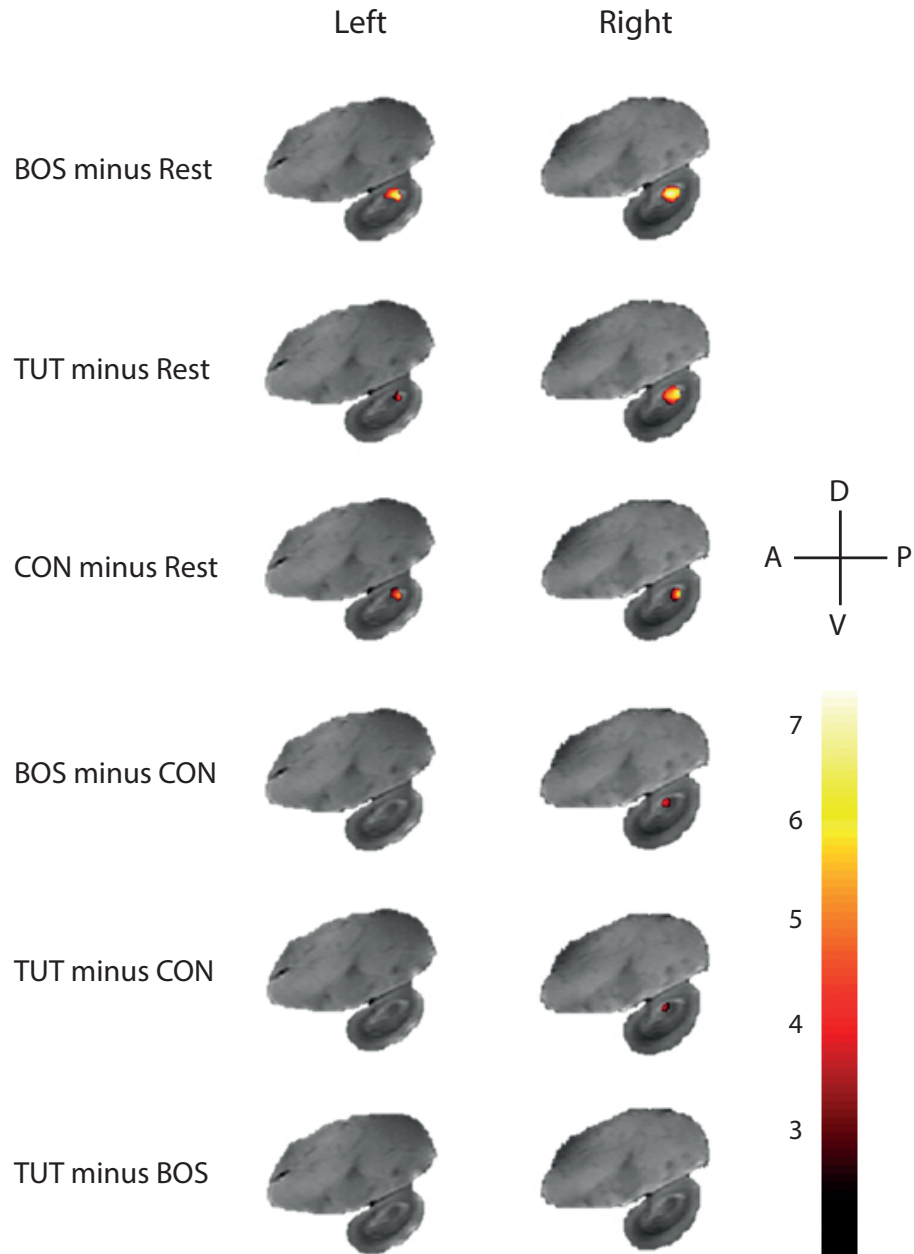
**Figure 3.4: Statistical maps of BOLD activation induced by the different stimuli in left and right MLd.** Results are superimposed on sagittal anatomical slices coming from the MRI zebra finch atlas. T-values are color-coded according to the scale displayed at the bottom of the figure. Note that the analysis was restricted to MLd and only voxels found to be significant (one-tailed t-test, $p < 0.05$, corrected at MLd level) are displayed. D: dorsal, V: ventral, A: anterior, P: posterior.
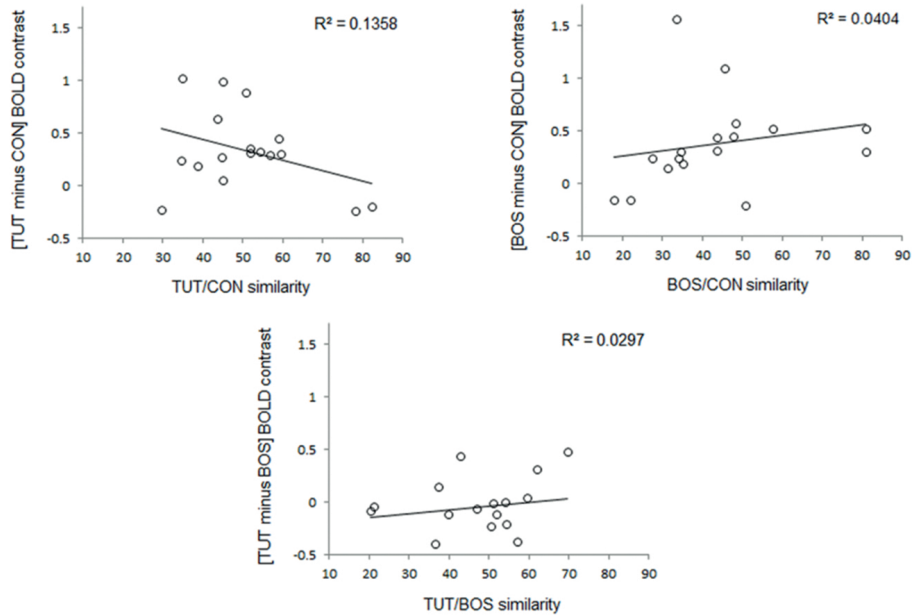
**Figure 3.5: Correlation between MRI signals and the acoustic similarity between the stimuli in right MLd.** The MRI signals (expressed in non-dimensional units) correspond to the mean amplitude estimate of the differential BOLD signals between TUT and CON (top left), BOS and CON (top right) and TUT and BOS (bottom). Positive values on the y axis indicate higher activations induced by the first stimulus of the comparison than the second one while negative values indicate higher activations induced by the second stimulus of the comparison than the first one. All correlations are statistically non-significant.

plitude of BOS and TUT selectivity and the age of birds. The two correlations were non-significant ([TUT minus CON]: $R^2 < 0.01$, $p = 0.80$, [BOS minus CON]: $R^2 < 0.01$, $p = 0.78$).

### 3.3.3 Brain responses in NCM

Left and right NCM were significantly positively activated by the three stimuli (Figure 3.7; Left NCM: [BOS minus Rest]: tmax $= 22.3$, $p < 0.0001$; [TUT minus Rest]: tmax $= 22.2$, $p < 0.0001$; [CON minus Rest]: tmax $= 22.4$, $p < 0.0001$; Right NCM: [BOS minus Rest]: tmax $= 32.2$, $p < 0.0001$; [TUT minus Rest]: tmax $= 33.9$, $p < 0.0001$; [CON minus Rest]: tmax $= 33.1$, $p < 0.0001$). We did not find any significant difference in term of BOLD response amplitude between the stimuli (Left NCM: Fmax $= 3.0$, $p = 0.88$; Right NCM: Fmax $= 4.4$, $p = 0.65$).

The lack of significant differential activation in NCM prevented us to test for potential correlation between differential activations and learning strength. Nevertheless, a correlation between [TUT minus Rest] and learning
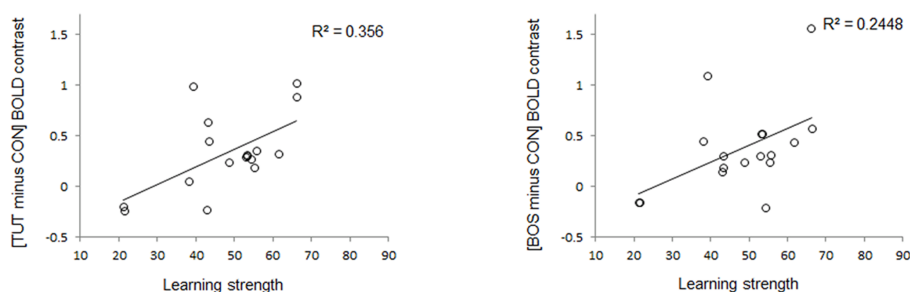
**Figure 3.6: Correlation of TUT (left) and BOS (right) selectivity with vocal learning strength in right MLd.** TUT and BOS selectivity are expressed as the mean amplitude estimate of the differential BOLD signals of [TUT minus CON], and [BOS minus CON], in non-dimensional units. Positive values on the y axis indicate a higher activation induced by TUT (or BOS) compared to CON while negative values indicate a higher activation induced by CON compared to TUT (or BOS). Both correlations are statistically significant.

strength could be expected in NCM based on earlier studies (Bolhuis et al., 2001, 2000; Terpstra et al., 2004). Such analysis failed to reveal any significant correlation (left NCM: $R^2$max $= 0.36$, $p = 0.15$, Right NCM: $R^2$max $= 0.09$, $p = 0.86$). However one can notice that the maximal correlation value measured in left NCM was of the same magnitude as the one measured between TUT selectivity and learning strength in right MLd ($R^2 = 0.36$ for both correlations). The big difference in terms of $p$-values is due to the correction for multiple tests applied in NCM (corrected/ uncorrected $p$-value = 0.15/0.006), which is directly related to the size of the investigated region. The correlation analyses performed on NCM were thus much less sensitive than the ones performed on MLd. Interestingly, a cluster of voxels in left NCM surviving the uncorrected p threshold of 0.05 was located in the posterior and lateral part of NCM (Figure 3.8), where Bolhuis and colleagues previously found a significant correlation between tutor song evoked gene expression and learning strength (Bolhuis et al., 2001, 2000; Terpstra et al., 2004).

Intrigued by this similitude, we further explored whether the correlation with learning strength was specific to tutor song or whether similar results could be found for BOS and CON evoked activations. These last analyses revealed no correlation of learning strength with [BOS minus Rest] and [CON minus Rest] (Figure 3.9, $R^2 < 0.14$; uncorrected $p$-values $> 0.14$), suggesting that as in Terpstra et al. (2004), the correlation was specific to the tutor song.

## 3.4    Discussion

The present study demonstrates selectivity for tutor song and bird's own song in right MLd, the main auditory midbrain nucleus. This selectivity was defined by a higher BOLD response induced by TUT and BOS than by CON.
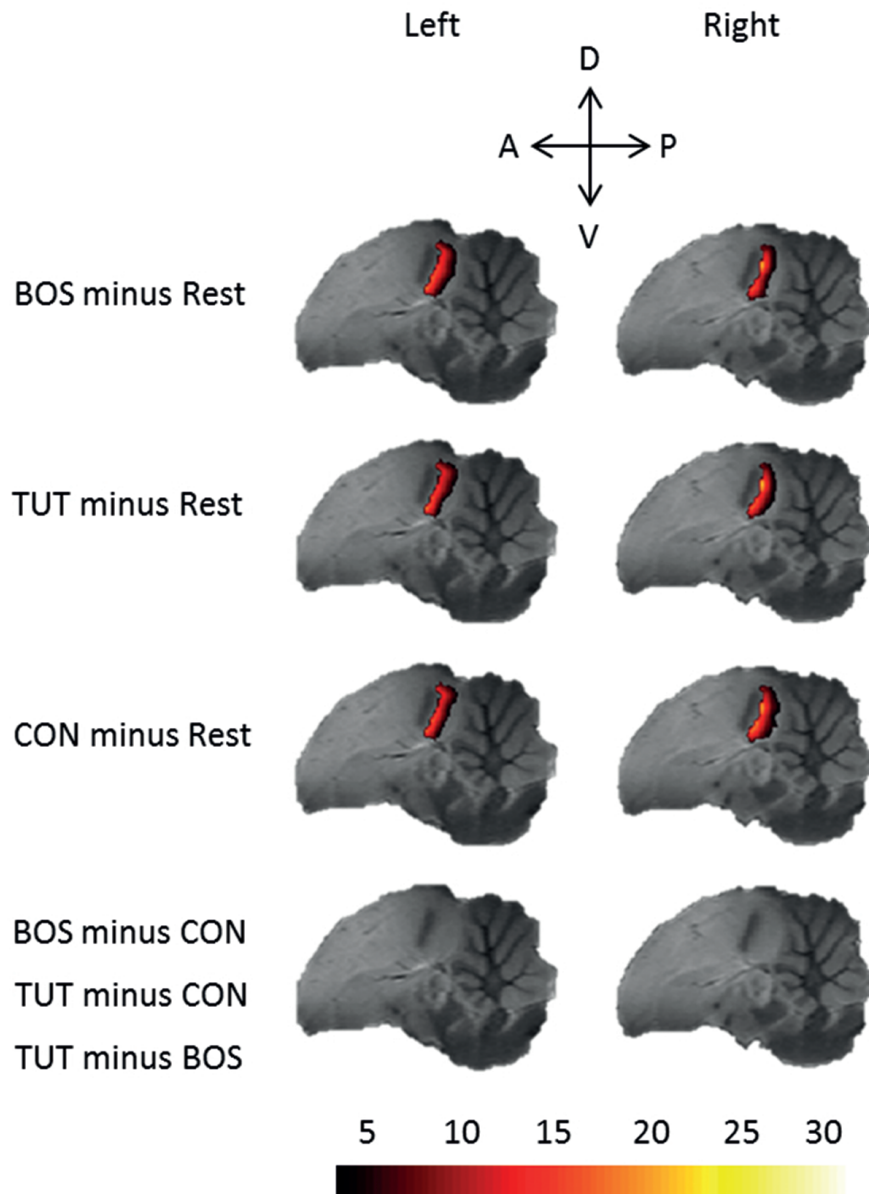
**Figure 3.7: Statistical maps of BOLD activation induced by the different stimuli in left and right NCM.** Results are superimposed on sagittal anatomical slices coming from the zebra finch MRI atlas. T-values are color-coded according to the scale displayed at the bottom of the figure. Note that in the figure other auditory regions (Field L and caudo-medial mesopallium) seem not activated only because the statistical analysis was restricted to NCM (for the whole activation pattern in the telencephalic auditory regions, see Fig. 3.2). Only significant voxels (one-tailed t-test, $p < 0.05$, FWE-corrected for NCM) are displayed. D: dorsal, V: ventral, A: anterior, P: posterior.
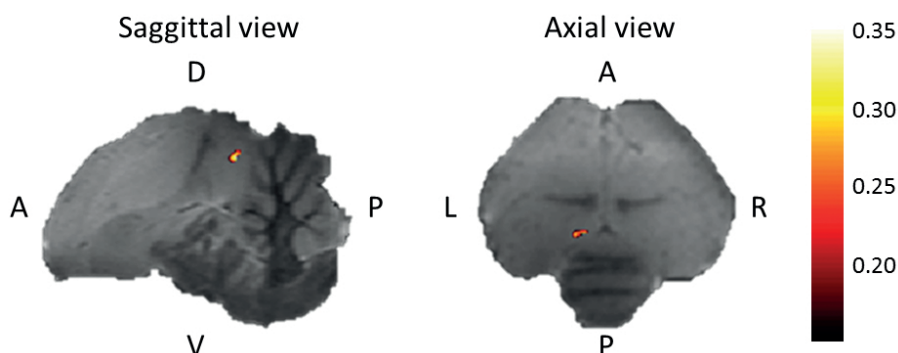
**Figure 3.8: Map of correlation between TUT minus Rest and vocal learning strength in left NCM.** Results are superimposed on sagittal and axial anatomical slices coming from the zebra finch MRI atlas and displayed at a threshold of $p < 0.05$ without correction for multiple tests. $R^2$ values are color-coded according to the scale displayed at the right side of the figure. D: dorsal, V: ventral, A: anterior, P: posterior; L: left; R: right.

The impact of acoustic features was controlled by using a conspecific song acoustically close to BOS and TUT and by a posteriori testing potential correlation between the strength of selective signals and the estimated amplitude of the residual acoustic differences between the stimuli. Such correlations were found not significant, ruling out the acoustic parameters as the main experimental factor responsible for the selectivity. This result rather suggests that it is the interaction between the acoustic features and the stimulus history which is responsible for the selectivity. The nature of the stimulus history responsible for the selectivity can be narrowed down since we used a familiar conspecific song as a control stimulus. The conspecific song came from a bird housed with the experimental bird after the end of the sensori-motor learning period (i.e. after 100 DPH), indicating that selective signals were induced by songs learned during the sensory-motor learning period.

Since the tutor song and the bird's own song are usually acoustically close, it has been suggested that responses to the tutor song might reflect sensitivity to the bird's own song (Yazaki-Sugiyama & Mooney, 2004). In the present study, BOS and TUT stimuli induced BOLD responses of similar amplitude. However, if the acoustic similarity was responsible for this lack of significant difference, similar BOLD responses should have been also found between BOS and CON since the acoustic similarity was not significantly different between each pair of stimuli. On the contrary, BOS and CON induced neural responses of significantly different amplitude. One would also expect the difference between BOS and TUT BOLD responses to be negatively correlated with the acoustic similarity between the two stimuli, which was not the case in the present study. Altogether, these results indicate that the right MLd is selective for both stimuli. BOLD fMRI signal reflects the activity of large populations of neurons. It is thus possible that different neuronal sub-populations are se-
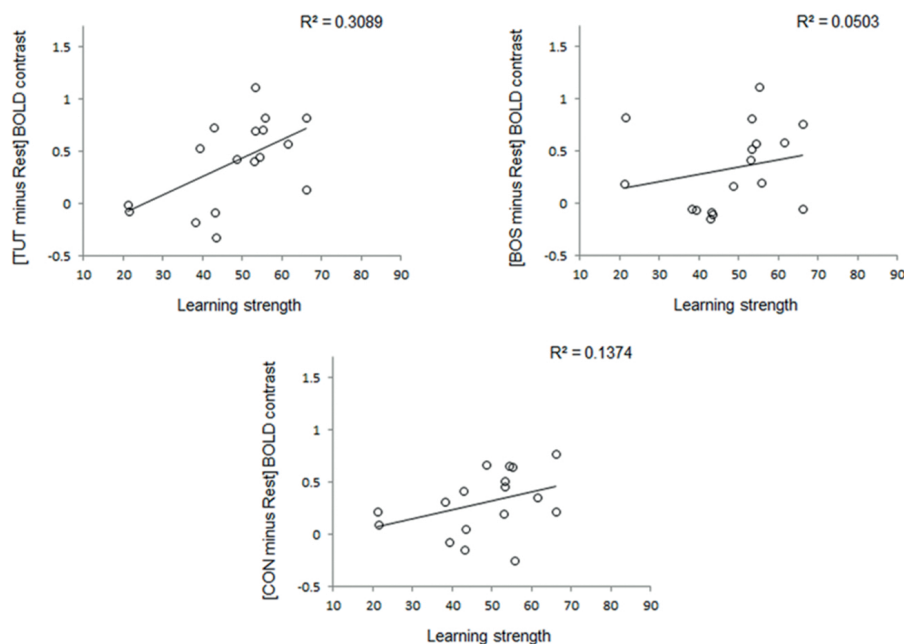
**Figure 3.9: Correlation of TUT, BOS and CON responsiveness with vocal learning strength in left NCM.** TUT, BOS and CON responsiveness are expressed as the mean amplitude estimates of the BOLD activations [TUT minus Rest], [BOS minus Rest] and [CON minus Rest], in non-dimensional units) in the left NCM cluster illustrated in Fig. 3.8. Note that the $R^2$ value in the left panel (0.3089) corresponds to the correlation value between learning strength and the [TUT minus Rest] signal averaged over the NCM cluster illustrated in Fig. 3.8 whereas the value reported in the text (0.36) corresponds to the correlation in the voxel where this correlation is the highest ($R^2$max). These two $R^2$ values are significantly different from 0. Correlation of BOS and CON responsiveness with learning strength are not significant.

lective for the bird's own song and the tutor song. Alternatively, the same neurons could be selective for the two types of stimuli, as it has been shown in few neurons of the anterior forebrain pathway (Solis & Doupe, 2000). The tutor song selectivity found in the right auditory midbrain indicates that a representation of the tutor song is still present in the adult brain. Since the tutor song is the song memorized by the experimental bird and later used to guide its vocal practice, the presence of selective responses which cannot be explained by acoustic differences between the stimuli strongly suggest that MLd is part of the neural substrates of tutor song memory. Reinforcing this interpretation, the strength of TUT selectivity was found to be positively correlated with the amount of song that the experimental birds copied from their tutor. This correlation suggests that birds that formed an accurate or well-consolidated memory of their tutor song later produced an accurate copy of this song.

BOS selectivity in right MLd constitutes an important replication of our

previous findings (Poirier et al., 2009). The present study demonstrates that this selectivity can be detected even when the conspecific song used as a control stimulus is acoustically close to the bird's own song. Birdsong is thought to be learned by trial and error. Detecting vocal errors supposes to identify the current state of the bird's own song via the auditory feedback, and then to compare it with the memorized tutor song. Bird's own song selective responses are thought to support these mechanisms (Prather & Mooney, 2004; Theunissen et al., 2004). Bird's own song selectivity in right MLd could thus reflect the identification of the bird's own song current state or the output of the comparison between the current song and the tutor song memory. The strength of bird's own song selectivity in MLd was found positively correlated with the amount of song experimental birds copied from their tutor. This result might suggest that bird's own song selectivity reflects the output of the comparison, the selective signal being stronger when the current song is found closer to the tutor song memory. Alternatively, this correlation could reflect the accuracy of bird's own song current state identification: indeed, an accurate bird's own song encoding is necessary to produce an accurate copy of the tutor song. Since tutor song selective responses were also found in the same nucleus, the subsequent comparison of the current bird's own song with the tutor song memory could then be made in MLd main efferent target, the auditory nucleus of the thalamus, and/or downstream, in the telencephalic auditory regions. This hypothesis is supported by recent evidence indicating that neurons in these thalamic and telencephalic regions increase their activity in response to feedback perturbations and thus could encode information about the quality of the bird's own song relative to the tutor song (Keller & Hahnloser, 2009; Lei & Mooney, 2010).

Numerous studies have pointed to another region of the ascending auditory pathway, NCM, to be involved in tutor song memory (Bolhuis & Gahr, 2006; Gobes & Bolhuis, 2007; Phan et al., 2006). One of these studies has shown that despite a similar amount of immediate early gene expression evoked by the tutor song, the bird's own song and a novel song in the lateral part of NCM of adult birds, only the activity evoked by the tutor song was positively correlated with the quality of tutor song imitation (Bolhuis et al., 2000). A similar trend was observed in the present fMRI study. In the ascending auditory pathway, MLd sends projection to the auditory nucleus of the thalamus called Ovoidalis, which projects to Field L at the telencephalic level. Field L then projects to NCM and the caudal mesopallium (CM) (Figure 3.10). Along this pathway, the information is considered to be encoded in a hierarchical way, neurons in NCM and CM being more complex than those in MLd (for a recent review, see Woolley, 2012). For instance MLd is known to respond to a wide variety of sounds, including conspecific and heterospecific songs but also tones and white noise while NCM mainly responds to conspecific songs. MLd neuronal responses are also more reliable, encoding precisely the spectro-temporal characteristics of the stimuli and are less context-dependent

than NCM responses.

While our results are consistent with recent evidence showing that MLd can encode the identity of individual songs (Schneider & Woolley, 2010) and that their activity can be modulated by early auditory experience (Woolley, Hauber, & Theunissen, 2010), the fact that tutor and bird's own song selectivity was found in the MLd of adult birds and not in NCM does not fit well with a hierarchical organization. We cannot rule out that the lack of selectivity in NCM is not due to the limited sensitivity of our experiment. Alternatively, the fact that the correlation of neural activity with learning strength was associated with selectivity for the tutor song in MLd but not in NCM suggests that the two regions play different roles putatively supported by different underlying mechanisms and different neural pathways. It has been recently demonstrated that the nucleus interface of the nidopallium (NIf) and HVC (used as a proper name), two pre-motor nuclei displaying bird's own song selective responses, play an important role in tutor song encoding (Roberts, Gobes, Murugan, Ölveczky, & Mooney, 2012). The nucleus ovoidalis is suspected to send projections to NIf (Wild, 2004), which project to HVC. MLd selective responses could thus reflect activity in this alternative pathway. Finally, the shelf of HVC sends projection to the area surrounding the nucleus robustus of the arcopallium (RA) which projects to Ovoidalis and MLd (Figure 3.10). Our results might thus reflect activity in these descending projections.

MLd tutor song and bird's own song selective signals described in the present study have been detected in anesthetized birds. A recent report indicates that tuning properties of MLd neurons are similar in awake and anesthetised individuals (Schumacher et al., 2011). Additionally, results of the present experiment in NCM constitute a replication of what have been found with another technique in awake birds (Terpstra et al., 2004), suggesting that anesthesia did not have a strong influence on the results. On the other hand, bird's own song selective responses in other forebrain regions have been found to be present when birds are anesthetised or asleep but to vanish when birds are alert (Cardin & Schmidt, 2004; Dave et al., 1998). Because these selective responses mimic spontaneous on-going activity occurring during sleep, they have been interpreted as reflecting off-line memory consolidation processes (Dave & Margoliash, 2000). Playback of tutor song during the day has also been found to induce in juvenile birds specific changes in bursting activity of neurons during the following night of sleep, suggesting again that memory consolidation processes took place during the night (Shank & Margoliash, 2009). Tutor song and bird's own song selective signals found in MLd might thus alternatively reflect such off-line memory consolidation processes. Either way (on-line or off-line mechanisms), the behavioural relevance of MLd selective signals in term of song learning is supported by the correlation found between the strength of the selectivity and the amount of song juvenile birds copied from their tutor.

Finally, bird's own song and tutor song selectivity was found in right but

not left MLd. Even if investigating the lateralization of the responses was beyond the scope of this study, these results comfort the right lateralization of bird's own song selective responses found in MLd in our previous study (Poirier et al., 2009). A recent study suggests that lateralization for conspecific song at the telencephalic level depends on auditory experience (Phan & Vicario, 2010). At the midbrain level, auditory experience has been shown to influence information coding and firing rate of MLd neurons (Woolley et al., 2010). Whether lateralization of MLd responses is also experience-dependent should be the object of further investigation.

To conclude, this study indicates that a memory trace of the vocalizations used as a model to guide vocal learning is present in the right auditory midbrain of adult songbirds. By showing that early sensory experience can generate long-lasting memories in a brainstem structure, it provides additional evidence to the growing body of research showing that that experience-dependent plasticity is not limited to cortical structures (Tzounopoulos & Kraus, 2009; Xiong et al., 2009). Recent studies indicate that the human auditory brainstem is involved in foreign language learning (Chandrasekaran, Kraus, & Wong, 2012; Song, Skoe, Wong, & Kraus, 2008) and training-based
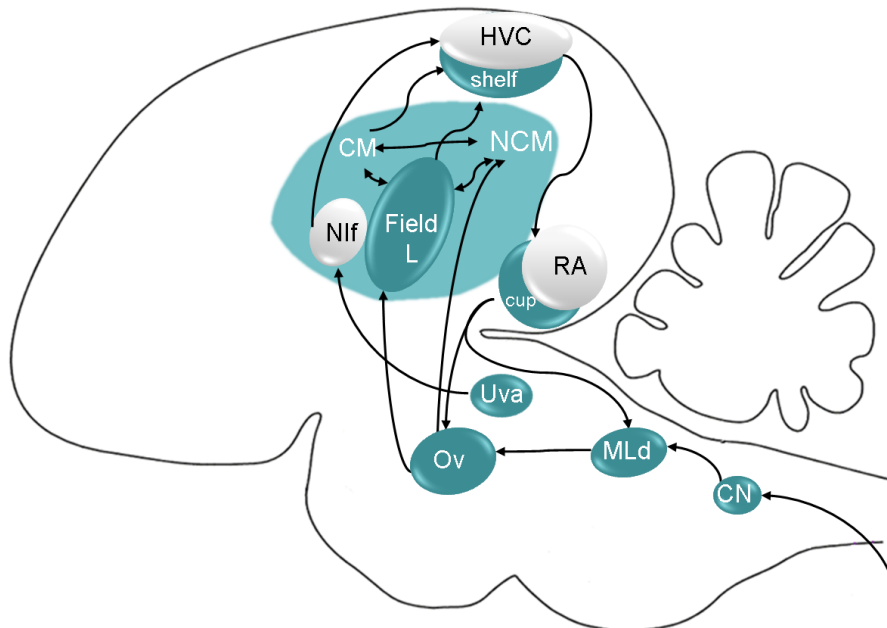


**Figure 3.10: Schematic representation of the songbird brain (parasagittal view).** The auditory regions are in blue and the vocal motor regions in grey. Only the main connections are represented. NIf: nucleus interface of the nidopallium; Ov: nucleus ovoidalis; RA: nucleus robustus of the arcopallium; Uva: nucleus uvaeformis; CN: cochlear nucleus.

improvement of speech hearing in noise (Song, Skoe, Banai, & Kraus, 2012) in adults. Since the organization of the auditory pathway at the sub-cortical level is well conserved among vertebrates, the involvement of the auditory midbrain in the auditory memory supporting vocal learning might be important for both avian and mammalian vocal learners.

## 3.5 Acknowledgments

CHAPTER 4

# Neural correlates of zebra finch tutor song perception during development: changes in selectivity and lateralization

*Birdsong learning and human speech acquisition share interesting similarities. Since both songbirds and humans need to be exposed to the vocalizations of their parents within a limited time window early in life, a sensitive period, in order to develop their own, songbirds present an excellent model for studying the sensitive period for human speech acquisition. During this sensitive period, perceptual sensitivity to adult vocalizations is important for learning in both species. Here we used functional MRI to investigate the development of conspecific song and tutor song selectivity in the auditory midbrain of juvenile zebra finches. Selective responses were previously identified in the auditory midbrain of adult birds (Poirier et al., 2009; Van der Kant et al., 2013). Our data show neural selectivity for conspecific song in both left and right auditory midbrain of female zebra finches early in development, while in male zebra finches tutor song selectivity is mainly present in the left auditory midbrain nucleus towards the end of the sensory phase of song learning. Moreover, the lateralization of tutor song selectivity changes towards adulthood. The strong tutor song selectivity near the end of the sensory period in males might reflect consolidation of a memory trace for the tutor's song. The difference in lateralization between juveniles and adults is reminiscent of the change in lateralization of human language in the course of development. While language in infants is largely bilateral, most human adults show a left lateralization of language.*

## 4.1   Introduction

Both juvenile songbirds and human infants learn their vocalizations by imitation and need to do so within a limited time window, called the sensitive period for vocal learning. Songbirds are therefore the animal model of choice to study vocal learning. Memories of adult vocalizations play an important role in this process, because vocal practice does not always take place in the vicinity of a tutor. Zebra finches memorize the song of their tutor during a sensory phase, which is followed by a sensorimotor phase, during which the birds match their vocalizations to the tutor's song (Eales, 1985). When reared in isolation from adult vocalizations, a memory will not be formed and neither songbirds nor humans will successfully develop their own vocalizations (Doupe & Kuhl, 1999; P. H. Price, 1979).

Several studies have proposed the auditory lobule (London & Clayton, 2008) and the caudomedial nidopallium (NCM) in particular (Bolhuis et al., 2001, 2000; Gobes, Zandbergen, & Bolhuis, 2010; Phan et al., 2006) as candidate regions for the neural substrates for tutor song memory in adult zebra finches. In juvenile zebra finches NCM has been demonstrated to show tutor song induced responses around 56 days of age (Gobes et al., 2010), but not before the start of the sensorimotor phase (Amin et al., 2007). In our previous study (Van der Kant et al., 2013) we showed tutor song selectivity in the right auditory midbrain nucleus (MLd) of adult male zebra finches.

The present study aims to longitudinally map tutor song perception throughout song development within-subject in order to show the dynamics of tutor song memory during song learning. We used in vivo Blood Oxygen Level Dependent (BOLD) functional Magnetic Resonance Imaging (fMRI) in juvenile male and female zebra finches at different stages of song learning in order to address tutor song selectivity in MLd. We compared the BOLD responses elicited by the birds tutor song and an unfamiliar conspecific song in each bird at 30, 40, 60 and 100 days post hatching (DPH). Additionally, we addressed conspecific song selectivity by comparing BOLD responses to unfamiliar conspecific and heterospecific song. Additional male data were obtained at ∼500 DPH for comparison with our previous adult results. Furthermore, we explored lateralization differences in tutor song selectivity over development.

Based on our previous study (Van der Kant et al., 2013) we expect males to show a gradual development of this neural selectivity for the tutor song throughout the sensory phase, exceeding adult levels. The enhanced selectivity is expected to drop towards adult levels at the end of the sensitive period, when it is no longer used for sensorimotor learning. Females do not use tutor song memory for motor learning, but have a preference for the tutor song in adulthood (Riebel, 2000). Female tutor song selectivity is therefore expected to be present at a lower level compared to adult males.
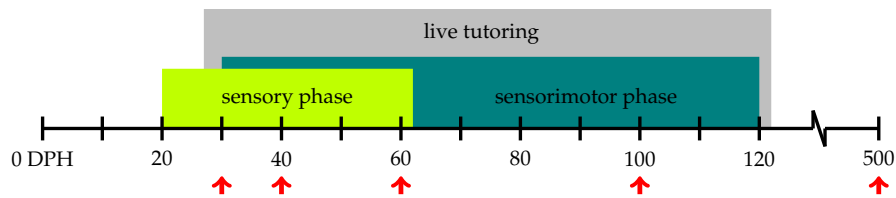
**Figure 4.1: Tutoring- and measuring paradigm.** fMRI data were acquired at 30, 40, 60 and 100 DPH for both sexes and additionally at 500 DPH for males (red arrows). Lime and teal bars show the sensory and sensorimotor phases, which partly overlap in zebra finches. The gray bar indicates the period during which the birds were individually housed with an adult tutor.

## 4.2 Materials and methods

### 4.2.1 Subjects

Nine juvenile zebra finches (Taeniopygia guttata) (5 males; 4 females) from the breeding colony of the Bio-Imaging Lab at the University of Antwerp (Belgium) were used for the present study.

The birds were raised by both parents in an individual cage from 0 to 7 DPH, by their mother from 8 DPH until they were fully fledged and able to feed independently (29 DPH) and were housed singly with an adult male tutor in a sound attenuated chamber from 30 to 120 DPH. This protocol is known to optimize song learning (Tchernichovski, Mitra, Lints, & Nottebohm, 2001). Following tutoring and after the last fMRI measurement, the male and female birds were housed in same-sex common aviaries.

All birds were maintained under a 12h light: 12h dark photoperiod throughout the experiment and had access to food, water and baths ad libitum. All experimental procedures were performed according to the Belgian laws on the protection and welfare of animals and were approved by the ethical committee of the University of Antwerp, Belgium (EC no. 2011/08).

### 4.2.2 fMRI procedure

The birds included in the present study received fMRI measurements at a number of time points during their song development (Figure 4.1). All nine birds were measured at the first four time points: 30, 40, 60 and 100 DPH. In order to match our present results with those from our last study which showed tutor song selectivity in the right MLd of adult males (Van der Kant et al., 2013), the male birds were also measured later in adulthood (around 500 DPH). Because one of the male birds was lost before the last measurement, data could only be obtained from four of the male birds at this time point.

**Stimuli**

For each experimental bird, three undirected songs were used as stimuli in the fMRI experiment: the tutor song (TUT), the song of an unfamiliar conspecific (CON) and a starling song (HET).

Stimuli were constructed using song material previously recorded from all tutors and colony birds. For each bird, the most representative motif (based on frequency of occurrence) was selected for use as a stimulus (Van der Kant et al., 2013). These motifs were concatenated into strings repeating the same motif, interleaved with silent periods. The duration of the silent periods was selected such that the total amount of song and silence was equal across all stimuli and mean intensity was normalized at 70dB SPL. An equalizer function was applied to the stimuli using WaveLab software (Steinberg, Germany) to correct for enhancement of frequencies between 2500 and 5000 Hz in the magnet bore (Poirier et al., 2010).

**Experimental setup**

During the experiment birds were continuously anesthetized with isoflurane (anesthesia induction, 3%; maintenance, 1,5%). The anesthetic gas was inhaled with a mixture of oxygen and nitrogen (at flow rates of 100 and 200 $cm^3$/min, respectively) through a mask over the birds beak. Body temperature was continuously monitored with an internal probe and maintained at 40°C by a feedback-controlled heating system (SA-Instruments, Stony Brook, NY, USA). The respiration rate was continuously monitored with a pneumatic sensor (SA-Instruments) positioned under the chest of the bird.

Auditory stimulation was delivered through dynamic loudspeakers (Visation, Germany) from which the magnets were removed, placed at both sides of the bird's head, close to the ears of the bird (Van Meir et al., 2005). Stimulus delivery was controlled by Presentation 0.76 software (Neurobehavioral Systems Inc., Albany, CA, USA).

**Procedures**

During fMRI acquisition each bird was presented with one TUT, one CON and one HET stimulus, pseudo-randomly repeated within a single experiment. Each bird received the same auditory stimuli at each time point during development. Stimuli were presented in an ON/OFF blocked design alternating 16s stimulation (ON blocks) and 16s rest periods (OFF blocks). During each block two images were acquired. Each stimulus type was presented 31 times, resulting in the acquisition of 62 images per stimulus and per subject. Three different stimulus orders were used, randomized between birds.

**fMRI data acquisition**

Blood Oxygen Level Dependent (BOLD) fMRI images were acquired using a 7T Bruker Pharmascan small animal MRI system (Bruker, Erlangen, Germany). The system and the RF coils used for the first four time points of the experiment have been described in a previous publication (Boumans et al., 2007). The last (500 DPH) time point was acquired with a RF transmit volume coil and a RF receive quadrature linear array mouse head surface coil (Bruker, Erlangen, Germany) using the same imaging parameters. For each bird, a time series of 382 $T_2$-weighted SE RARE images was acquired in the sagittal plane, covering the whole brain (15 slices, 750 $\mu$m thick, 50 $\mu$m gap, 250 $\mu$m$^2$ in-plane resolution, Effective Echo time (TEeff) / Repetition time (TR): 60/2000 ms, RARE factor: 8, Field of View: 16 x 16 mm), which resulted in a good signal in bilateral MLd. The first 10 images from each time series were considered dummy images to allow the bird to get accustomed to the scanner noise before the presentation of the stimuli was started. Following the fMRI acquisition, a high-resolution structural 3D SE RARE volume (voxel size 125 $\mu$m$^3$; TEeff/TR: 60/2000 ms; RARE factor: 8; Field of View: 16 x 16 mm) was acquired for each bird to allow for spatial registration with the zebra finch atlas (Poirier et al., 2008).

**fMRI data analysis**

Data were preprocessed and analyzed using Statistical Parametric Mapping software (SPM8, Wellcome Trust Centre for Neuroimaging, London, UK). The first 10 (dummy) images from each time series were excluded from the analysis. Each fMRI time series was realigned using a least-squares procedure and a six-parameter rigid-body spatial transformation. Head movement did not exceed 500 $\mu$m for any bird. To enable localization of activations in specific brain regions, the high-resolution 3D image of each subject was spatially registered to the MRI atlas of the zebra finch brain (Poirier et al., 2008). The resulting spatial transformation matrix was then applied to the functional time series and re-sampled to obtain a resolution of 125 x 125 x 400 $\mu$m. These steps resulted in a good correspondence between anatomical areas in the atlas and the fMRI data. Finally, the fMRI images were smoothed with a Gaussian kernel of 500 $\mu$m FWHM in-plane and 800 $\mu$m FWHM in the slice direction.

For each bird the BOLD response was modeled as a box-car function convolved with a canonical hemodynamic response function using the General Linear Model and a classical restricted maximum likelihood algorithm. Subject-level fixed-effect analyses were then performed for the comparison of each stimulus with the rest condition. To model the possible effects of head motion (which is not convolved with the hemodynamic response), motion (realignment) parameters were included in the model as regressors for each subject.

The contrasts obtained in the subject-level analysis were entered into a

group-level random effects analysis. Based on earlier results (Van der Kant et al., 2013), the group analysis was restricted to right and left MLd (the auditory midbrain nuclei), defined as Regions of Interest (ROI's). For the anatomical location of these ROI's, see Figure 3.3 on page 54. Developmental changes in the BOLD response were expected in MLd based on TUT selectivity previously measured in MLd (Van der Kant et al., 2013). Right and left MLd could be clearly identified and delineated on the zebra finch atlas (Poirier et al., 2008).

Voxel-based statistical differences between stimulus-related BOLD signal changes were assessed using a one-way repeated measures ANOVA followed by post-hoc uni-lateral paired t-tests in the two regions of interest. Differences between the sexes following from interactions were assessed with two-sample t-tests on the contrasts of interest. Only activations with a volume of at least 5 voxels were considered. P-values were adjusted for the number of independent tests performed using Family Wise Error correction for each anatomical Region of Interest (ROI), based on Random Field Theory (Worsley et al., 1996). When an ROI is reported to contain "no clusters", no clusters are found in the uncorrected data with a threshold of $p = 0.05$. Lateralization indices were used to assess laterality of the contrasts of interest. These indexes were computed between left and right MLd in a threshold-free manner using the "AveLI" SPM add-on (Matsuo, Chen, & Tseng, 2012), freely available at http://mrilab.mc.ntu.edu.tw/aveli/). This add-on computes the laterality index based on the sums of positive T-values within a selected ROI (in this case MLd) weighted for the number of voxels with positive T-values within that ROI. The LI represents the portion of the signal that is found either left or right with -1 representing 100% right lateralization and 1 representing 100% left lateralization.

## 4.3 Results

### 4.3.1 Tutor song selectivity

**Tutor song selectivity independent of sex**

Group analyses with males and females pooled showed that bilateral MLd was selectively activated by the tutor song compared to a novel conspecific song only at the age of 60 DPH (60 DPH [TUT-CON] in left MLd: mean sum T (the mean sum of positive t-values within the ROI) = 147.8, $sd = 103.3$, $p = 0.005$; [TUT-CON] in right MLd: mean sum T = 140.5, $sd = 169.5$, $p = 0.004$). However, no interaction of the TUT-CON contrast with age and sex was found in MLd. NCM did not show tutor song selectivity at any of the time points with males and females pooled and in analyses per sex and was thus excluded from further analysis. Because tutor song selectivity was expected to be more pronounced in male birds, a t-test was used to assess sex differences in the TUT-CON contrast at this age. This analysis revealed that the tutor song

selective response in left MLd was significantly stronger in male compared to female birds at 60 DPH (Males-Females 60 DPH: sum T at group level = 185.0, $p$ = 0.019). Based on this result, we further analyzed tutor song selectivity separately in males and females, in order to accurately describe the changes in tutor song selectivity throughout development for males and females.

**Tutor song selectivity in male MLd**

In male zebra finches, a trend towards a [TUT-CON $*$ time point] interaction was found in left MLd (sum F at group level = 321.9, $p$ = 0.081). Further analyses per time point showed that no tutor song selective response could be found in male birds at the early stages of song learning. At 30 and 40 DPH no tutor song selective clusters were present in the left MLd or in the right MLd (Figure 4.2). As also shown by figure 4.2, the perception of the individual stimuli does not show the same pattern at the ages of 30 and 40 DPH. While at 30 DPH the activation in left as well as right MLd is absent or only minor for all three stimuli (30 DPH: no clusters for [TUT-Rest] and [CON-Rest] in left and right MLd; [HET-Rest] in Left MLd: mean sum T = 58.9, $sd$ = 36.1, $p$ = 0.22; [HET-Rest] in Right MLd: mean sum T = 65.6, $sd$ = 67.3, $p$ = 0.09), left MLd does get significantly activated by TUT and CON as opposed to HET at 40 DPH ([TUT-Rest] at 40 DPH: mean sum T = 177.0, $sd$ = 85.2, $p$ = 0.012; [CON-Rest] at 40 DPH: mean sum T = 160.0, $sd$ = 65.2, $p$ = 0.008; [HET-Rest] at 40 DPH: no clusters). There is, however, no tutor song selective activation.

At 60 DPH, when the sensory phase reaches its end, male zebra finches show a strong tutor song selective response in the left auditory midbrain (MLd) ([TUT-CON] in left MLd: mean sum T = 207.3, $sd$ = 101.4, $p$ < 0.001). A post-hoc analysis of the individual stimuli showed that at this time point the activation in response to the tutor song remained high ([TUT-Rest] in left MLd: mean sum T = 145.7, $sd$ = 123.6, $p$ = 0.008), while in left MLd no activation was found for conspecific or heterospecific song. A TUT-CON cluster ([TUT-CON]: mean sum T = 207.5, $sd$ = 207.6, $p$ = 0.002) was also found in male right MLd at 60 DPH. However, due to the absence of a significant difference between the tutor song and heterospecific song related activity within

| | | TUT-CON | | CON-HET | | TUT-HET | |
|---|---|---|---|---|---|---|---|
| | DPH | T(sd) | $P$ | T(sd) | $P$ | T(sd) | $P$ |
| Left MLd | 60 | 207(101)* | 0.001 | | | | |
| | 100 | 223(282) | 0.11 | | | | |
| Right MLd | 500 | 43(31) | 0.081 | 87(23)* | 0.011 | 119(54)* | 0.001 |

**Table 4.1: Sums of T-values and $P$-values for male birds.** Values as reported in section 4.3. For contrasts for which male birds did not show a significant BOLD response ($p$ < 0.05), values are not reported. All values are FWE-corrected at the ROI level and only clusters which contain at least 5 voxels are reported
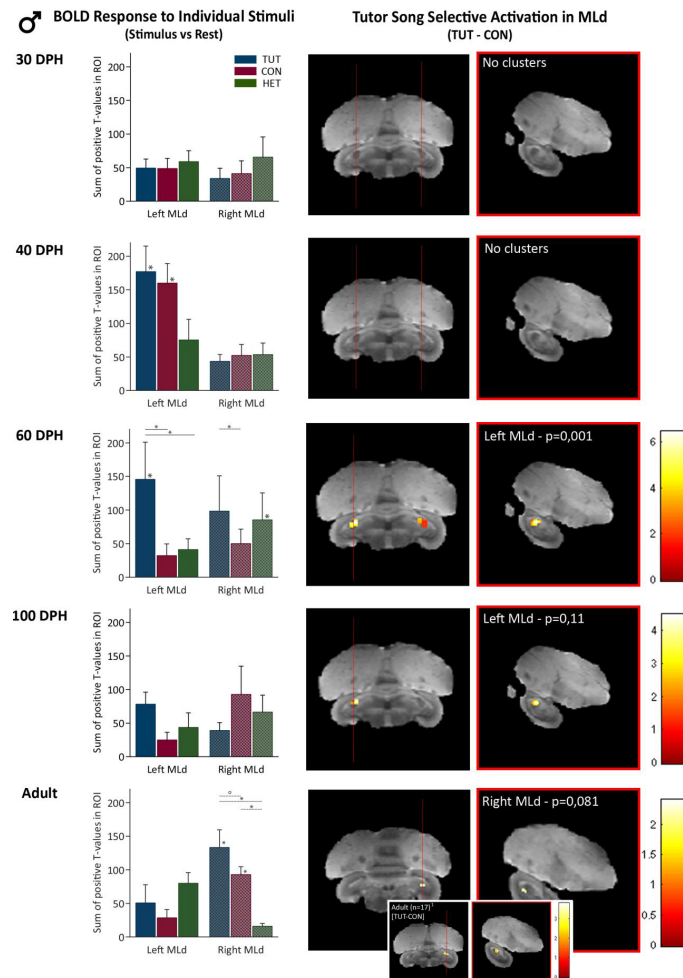
**Figure 4.2: LEFT: Mean BOLD response of Males to the individual stimuli (TUT, CON, HET).** The sum of positive T-values within each ROI (Left MLd / Right MLd) is used as a measure of BOLD activity. The zero level corresponds to the mean activation level during rest periods (exposure to scanner noise). Stars on top of the bars indicate statistically significant difference between stimuli vs. rest. Horizontal lines above the graphs indicate statistical significant differences between responses to the individual stimuli within the ROI ($^{\circ}p < 0.10$; $*p < 0.05$, corrected for multiple comparisons). **RIGHT: Male tutor song selectivity in MLd.** Only voxels showing a significant difference (threshold: $p < 0.05$) in BOLD response to TUT compared to CON in left and right MLd are displayed. The specified $p$-values in the sagittal images are FWE-corrected for the anatomical ROI (Left MLd or Right MLd) and only clusters containing at least 5 voxels are included. T-values are color-coded according to the scale displayed on the right side of each panel.

the cluster ([TUT-HET]: no clusters), this activation could not be confirmed as selective for the tutor song. Although right MLd was significantly more activated by tutor song compared to conspecific song, a large portion of the activation in right MLd at this age was induced by the heterospecific song ([TUT-Rest]: mean sum T $= 98.4$, $sd = 117.7$, $p = 0.188$; [CON-Rest]: no clusters; [HET-Rest]: mean sum T $= 85.3$, $sd = 89.7$, $p = 0.018$).

At 100 DPH, around the time of song crystallization, the tutor song selective response no longer reached significance in the males left MLd ([TUT-CON]: mean sum T $= 223.0$, $sd = 281.7$, $p = 0.11$; see figure 4.2). Furthermore, no tutor song selectivity was found in right MLd at this age (no [TUT-CON] clusters). Analysis of the individual stimuli in left MLd showed a trend for tutor song related activation, but left MLd was not activated by the other stimuli (Left MLd: [TUT-Rest]: mean sum T $= 78.2$, $sd = 40.0$, $p = 0.09$; [CON-Rest]: no clusters; [HET-Rest]: no clusters).

Earlier results showed adult male tutor song selectivity in right MLd (Van der Kant et al., 2013). In order to test whether the tutor song selective response in left MLd could be attributed to the age of the birds, male tutor song selectivity was additionally assessed at 500 DPH. At this age, the four remaining males showed a tutor song selective cluster in right MLd, which did not survive correction for multiple comparisons, but showed a trend (500 DPH [TUT-CON] : mean sum T $= 42.8$, $sd = 30.9$, $p = 0.081$; [TUT-Rest]: mean sum T $= 133.2$, $sd = 52.5$, $p = 0.004$; [CON-Rest]: mean sum T $= 92.8$, $sd = 23.5$, $p = 0.01$; [HET-Rest]: mean sum T $= 16.1$, $sd = 7.7$, $p = 0.004$).

**Lateralization of male tutor song selectivity**

Due to the interference of heterospecific song related activity in right MLd at 60 DPH, when tutor song selectivity is found to be prominent in males, a direct assessment of [TUT-CON] lateralization during song learning would not be informative. Instead, assessment of the activation induced by the individual stimuli does reveal a pattern throughout development.

In left MLd we see an initial increase of activation in response to both TUT and CON at 40 DPH. At the age of 60 DPH the CON related activation shows a sharp drop, leaving left MLd purely tutor song selective at this age. Later in development, the activation in response to TUT also drops, leaving left MLd selective for neither TUT nor CON. Activation in response to HET shows relatively minor changes throughout development in left MLd. In contrast, no large peaks in TUT or CON related activation are found in right MLd before adulthood.

At 500 DPH, TUT related activation shows a peak in right MLd, which is not found in left MLd. However, lateralization of the tutor song selective signal was highly variable at this age (mean LI [TUT-CON] $= -0.13$, $sd = 0.8$, $p = 0.36$). Individual analysis of the subjects showed that tutor song selectivity was highest in right MLd in three out of four subjects (LI's [TUT-CON]: -0.02, -0.99 and -0.45), while in the fourth subject, tutor song selectivity was highly

left lateralized (LI [TUT-CON] = 0.92).

**No significant tutor song selectivity in female MLd**

In females, no interaction between tutor song selectivity and time point was found in either right or left MLd. Like the males, the females did not show TUT selectivity at early age in left Mld ([TUT-CON] at 30 DPH: mean sum T = 47.3, $sd = 42.6$, $p = 0.28$; [TUT-CON] at 40 DPH: no clusters) nor in right Mld ([TUT-CON] at 30 DPH: no clusters; [TUT-CON] at 40 DPH: no clusters). In contrast to the males, however, tutor song did not evoke significant differential activation in left or right MLd of female zebra finches at 60 DPH, when compared to conspecific song ([TUT-CON] at 60 DPH in left MLd: mean sum T = 73.5, $sd = 38.8$, $p = 0.19$; in right MLd: no clusters). Analysis of the individual stimuli at this age confirms that none of the stimuli significantly activates left or right MLd in females ([TUT-Rest]; [CON-Rest]; [HET-Rest]: no clusters in left or right MLd). At 100 DPH no tutor song selectivity was found in MLd of female zebra finches ([TUT-CON]: no clusters in right or left MLd).

### 4.3.2 Conspecific song selectivity

Neural selectivity for conspecific song was assessed by comparing the BOLD activity evoked by a novel conspecific to that evoked by a novel heterospecific (European starling) song. The familiarity of both songs might increase with each measurement due to hearing the song under anesthesia, but the amount of cumulative exposure is the same for both songs throughout the experiment.

Selectivity for conspecific song as shown by the [CON-HET] contrast was found to be present in left MLd with sexes and time points pooled. Conspecific song selectivity showed a significant interaction between sex and time point in left MLd (sum F on group level = 355.8, $p = 0.018$). Therefore, conspecific song selectivity was further investigated in males and females separately.

**Conspecific song selectivity in males**

Although analysis of the individual stimuli shows a large peak of activation for TUT and CON in left MLd at 40 DPH in males (Left MLd at 40 DPH: [TUT-Rest]: mean sum T = 177.0, $sd = 85.2$, $p = 0.012$; [CON-Rest]: mean sum T = 160.0, $sd = 65.2$, $p = 0.008$; [HET-Rest]:no clusters), there is no significant conspecific song selectivity at this age (Left MLd: [CON-HET]: mean sum T = 99.3, $sd = 51.2$, $p = 0.13$; [TUT-HET]: mean sum T = 115.9, $sd = 58.3$, $p = 0.15$).

In adulthood, at 500 DPH, the right auditory midbrain is selectively activated by conspecific over heterospecific song (Right MLd [CON-HET]: mean sum T = 86.8, $sd = 23.0$, $p = 0.011$). This is confirmed by the tutor song ([Right MLd [TUT-HET]: mean sum T = 119.3, $sd = 54.2$, $p < 0.001$). There were no significant differences between conspecific and heterospecific song activations

in males at 30, 60, 100 and 500 DPH in left MLd and at any juvenile time points in right MLd.

**Conspecific song selectivity in females**

| | DPH | TUT-CON | | CON-HET | | TUT-HET | |
|---|---|---|---|---|---|---|---|
| | | T(sd) | $P$ | T(sd) | $P$ | T(sd) | $P$ |
| Left MLd | 30 | | | 178(91)* | 0.001 | 178(67)* | 0.001 |
| | 60 | 74(39) | 0.19 | | | | |

**Table 4.2: Sums of T-values and $P$-values for female birds.** Values as reported in section 4.3. For contrasts for which female birds did not show a significant BOLD response ($p < 0.05$), values are not reported. All values are FWE-corrected at the ROI level and only clusters which contain at least 5 voxels are reported

Analyses per time point demonstrated that MLd is only selective for conspecific song in females at 30 DPH, when they have spent only limited time with their tutor ([CON-HET] in left MLd at 30 DPH: mean sum T = 178.3, $sd = 90.8$, $p < 0.001$) (Figure 4.3). This selectivity was shown only in the left MLd. In right MLd none of the stimuli induce a significant activation at this age. After 30 days of age, no significant selectivity for conspecific song was found in left MLd ([CON-HET] at 40 DPH: mean sum T = 66.9, $sd = 71.5$, $p = 0.248$; [CON-HET] at 60 DPH: no clusters; [CON-HET] at 100 DPH: mean sum T = 71.1, $sd = 27.2$, $p = 0.238$).

In order to further investigate whether the selectivity at 30 DPH originated from hearing a novel conspecific rather than selectivity for conspecific song in general, we also assessed the differences in BOLD response elicited by the tutor song, which is familiar to the bird, compared to a heterospecific song. This analysis also showed significant differential activation in the left MLd (Left MLd [TUT-HET] at 30 DPH: mean sum T = 178.2, $sd = 66.7$, $p < 0.001$) (Figure 4.3). At 30 DPH left MLd is indeed significantly activated by tutor song as well as conspecific song (Left MLd [TUT-Rest]: mean sum T = 108.5, $sd = 60.0$, $p = 0.01$; [CON-Rest]: mean sum T = 107.6, $sd = 93.9$, $p = 0.049$), while showing no activation in response to heterospecific song (Left MLd [HET-Rest]: no clusters).

## 4.4  Discussion

The present study for the first time uses a longitudinal fMRI approach to assess song selectivity in juvenile zebra finches. This approach enabled us to follow the same group of birds throughout song learning and address changes in neural selectivity during development. In the present study we showed that the left MLd of juvenile male zebra finches shows strong selectivity for their
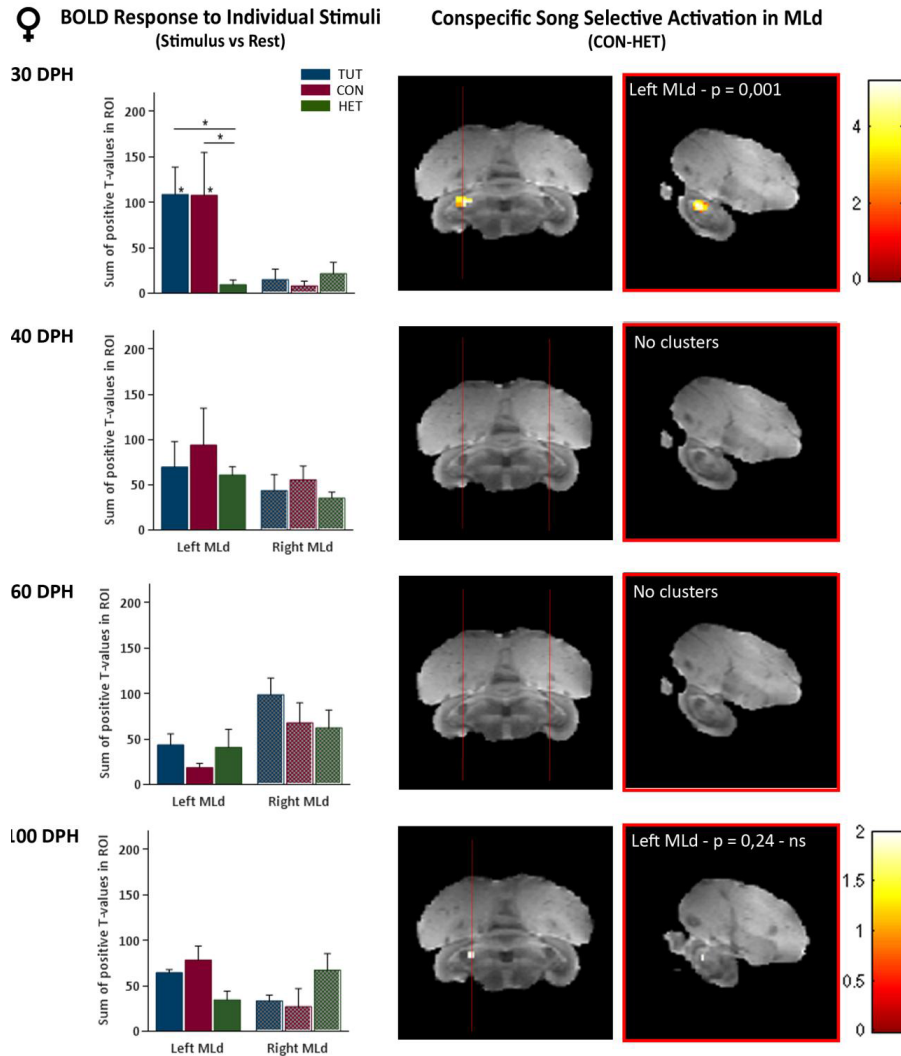
**Figure 4.3: LEFT: Mean BOLD response of Females to the individual stimuli (TUT, CON, HET).**
The sum of positive T-values within each ROI (Left MLd / Right MLd) is used as a measure of
BOLD activity. The zero level corresponds to the mean activation level during rest periods (ex-
posure to scanner noise). Stars on top of the bars indicate statistically significant difference be-
tween stimuli vs. rest. Horizontal lines above the graphs indicate statistical significant differences
between responses to the individual stimuli within the ROI ($p < 0.05$, corrected for multiple
comparisons). **RIGHT: Female conspecific song selectivity in MLd.** Only voxels showing a sig-
nificant difference (threshold: $p < 0.05$) in BOLD response to TUT compared to CON in left and
right MLd are displayed. The specified $p$-values in the sagittal images are FWE-corrected for the
anatomical ROI (Left MLd or Right MLd) and only clusters containing at least 5 voxels are in-
cluded. T-values are color coded according to the scale displayed on the right side of each panel.

tutor song at 60 DPH. Moreover, the left MLd of female zebra finches shows conspecific song selectivity over heterospecific song as early as 30 DPH. At this time, when they have only been exposed to their tutor for a limited time span, the female birds do not show selectivity for the tutor song over other conspecific song.

Tutor song selectivity was found in the male left MLd at 60 DPH. At this age, the sensory period of song learning approaches its end in zebra finches (Slater, Eales, & Clayton, 1988). Analysis of the activation in terms of the individual stimuli showed that this strong selectivity is caused by a combination of factors. Tutor song related activation in left MLd stays high compared to the 40 DPH time point, while conspecific song related activity diminishes. This might relate to the possibility of juveniles including elements from a second tutor in their song, which does still exist at 40 DPH, but is decreased when the birds reaches the age of 60-65 days, indicating that the sensory phase is ending (Eales, 1985). Therefore, when hearing a novel conspecific bird, the auditory system will not be highly activated anymore at this age. However, tutor song is still very important at the age of 60 DPH, because the bird is now in a stage of singing plastic song and needs to compare this song to the tutor song continuously. Therefore the memory of the tutor song needs to be a strong one. Moreover, sensory learning has reached its peak at this age and the memory of the tutor song formed during the sensory phase will be at its strongest. The finding that tutor song selectivity in left MLd cannot be found at 100 DPH indicates that this nucleus is mainly selective for tutor song during learning and is partly lost after song crystallization. Indeed, in an earlier fMRI study in adult male zebra finches (Van der Kant et al., 2013) no tutor song selectivity was found in left MLd.

In the abovementioned adult study, male zebra finches showed tutor song selectivity in the right auditory midbrain nucleus of adult male zebra finches. Although the last time point (500 DPH) of our longitudinal study contains only four males, it does show a trend towards tutor song selectivity in right MLd, which is consistent with our previous study (Van der Kant et al., 2013), showing significant tutor song selectivity only in right MLd, but no significant lateralization. Individual analyses of the four male birds also measured in later adulthood confirmed that three of these birds showed a left-to-right shift of tutor song selectivity between the ages of 60 DPH and 500 DPH. Although these data are too sparse to draw any conclusions towards a left-to-right shift during song development in male zebra finches, they seem to be in accordance our previous data, which showed tutor song selectivity in the right MLd of adult male zebra finches. Moreover, a similar change in lateralization of song perception over development was recently found in the zebra finch NCM using ZENK expression (Moorman et al., 2012). Together, these data might suggest that there is a broad change in lateralization in the brain regions which play a role in song memorization. This is the case in humans, where adult left lateralization for language is not yet present in infants, but

emerges during development (for a review, see Minagawa-Kawai, Cristià, & Dupoux, 2011).

If there is indeed a left-right lateralization shift, it might suggest that left MLd tutor song selectivity found during development and right MLd tutor song selectivity in adulthood represent two different processes, possibly related to song learning and song maintenance respectively. In order to investigate this issue and uncover the mechanisms behind the developmental changes, a longitudinal study with a larger group of males is ongoing in the Bio-Imaging Lab.

Although the effect was only significant in males, at 60 DPH the left MLd of females was also activated by the tutor song to a larger extent than the neighboring time points (40 and 100 DPH). This might indicate that the preference for tutor song found earlier in females (Clayton, 1988; Riebel, 2000) also has a neural correlate during development and that the process of sensory learning, although perhaps stronger in males, might be similar in males and females. The timing of the peak of tutor song related activity, which is similar in males and females, suggests that, although females do not learn to sing, they do show signs of a sensory period at the neural level. Female zebra finches showed highly significant selectivity for conspecific songs in the left auditory midbrain nucleus at the age of 30 DPH. At this time, left MLd is significantly activated by the tutor song and an unfamiliar conspecific song, while showing no activation in response to heterospecific song. This early neural sensitivity to conspecific song in female birds might contribute to the selection of quality mates in adulthood (Lauay, Gerlach, Adkins-Regan, & DeVoogd, 2004).

The CON selective activation in left MLd of juvenile female zebra finches is also in line with selectivity for conspecific song in left MLd found earlier in an auditory fMRI study in adult males (Poirier et al., 2009). However, in the present study this early selectivity for conspecific songs was not found in male birds. This is surprising, because male birds would need to distinguish a possible model for song learning from other sounds. Moreover, it has been shown that zebra finches can behaviorally distinguish conspecific from heterospecific songs very early in life and this has even been proposed to be an innate ability (Braaten & Reynolds, 1999). We cannot exclude the possibility that a neural selectivity for heterospecific song is present in the midbrain at an early age the developmental fMRI study, which is ongoing and includes a larger group of birds might be able to address this issue.

A selective response for conspecific songs could be detected in the male right MLd in adulthood, at 500 DPH. However, at this age the birds were no longer kept in isolation and had been exposed to other conspecifics. The familiarity of the conspecific song played during the experiment had not changed at this time, because the conspecifics to which the birds were exposed during scanning were either no longer present in the laboratory or housed in a different aviary. The selectivity at 500 DPH could therefore not be due to recognition of the song from the bird's daily environment. A more likely explanation

is that social interaction in the aviary renders the male birds more sensitive to territorial behavior and thus to relatively novel conspecific song.

Together with our earlier data (Van der Kant et al., 2013), the present data show that the neural substrates of tutor song memory in male zebra finches during song learning and in adulthood show a different lateralization. These findings are in accordance with both songbird and human data on laterality changes during vocal development, suggesting that lateralization changes during birdsong learning are not restricted to MLd, but might occur throughout the auditory system.

# Part II

# Neural correlates of artificial grammar learning in human adults

In the first part of this thesis, the songbird model was employed in order to study the neural correlates of vocal learning in a species that shows some interesting behavioral similarities to humans with regard to the development of their species-specific vocalizations. The data collected in the songbird studies have shown that the brain selectively processes the song that is learned from the tutor early in life. Furthermore, these studies have shown that this selectivity develops and changes during the song learning process and that its strength is related to learning success.

In the remainder of this thesis, we will explore whether the neural correlates of human language learning share some of the characteristics found for birdsong learning. If the neural processes undelying birdsong learning and human language learning show similarities, this would strengthen the songbird model for vocal learning. For ethical reasons, humans cannot be isolated from speech input at any time during or after the sensitive period for language acquisition, which means that the full extent of their language knowledge cannot be controlled. Furthermore, human language is highly complex and learning it requires many levels of analysis. This means that "what is learned" is too extensive to study in a single experiment. However, studies of language deprived children have shown that the acquisition of syntactic structure is most impaired when humans start to learn language after the sensitive period for language acquistion.

Because the acquisition of syntactic structure is such a central part of human language learning, we employ an artificial grammar learning paradigm to simulate and study grammar learning in human adults. This study serves as a basis for further studies on artificial grammar learning, which will be focused on children and compare adult learners to children in order to measure the impact of the sensitive period on the neural correlates of human language acquisition. In the study described in Chapter 5 the learning process is studied on-line using auditory fMRI and related to learning outcomes. Furthermore, structural and functional connectivity related to artificial grammar learning are explored in Chapter 6.

CHAPTER 5

# Neural correlates of individual differences in non-adjacent dependency learning in human adults

*How does the brain reflect the learning of a new language? In previous studies, the effects of language learning on the state of the brain have been investigated, but mostly off-line (after learning). In the present study, individual differences in brain function were studied during auditory artificial grammar learning. The neural correlates of individual differences in adult language learning were investigated in a functional Magnetic Resonance Imaging (fMRI) study using an auditory artificial grammar learning paradigm. FMRI data from 20 adults were collected while they were presented with an artificial grammar containing non-adjacent dependencies and a control language lacking these dependencies. After exposure participants were tested on learning and generalization of the artificial grammar. Our data showed a correlation between the differential neural activation in response to the grammar containing non-adjacent dependencies in the Left Inferior Frontal Gyrus (LIFG) and the bilateral Superior Temporal Gyri and Insulae and the level at which individual subjects showed sensitivity to the non-adjacent dependencies. Furthermore, during the grammatical judgment task, LIFG showed a differential activation in response to correctly versus incorrectly judged violations. These results indicate that individual differences in learning success are reflected in neural activation that develops early in the learning process and is largely dependent on sensitivity to violations.*

## 5.1 Introduction

Both during first and second language acquisition learners are confronted with speech streams from which they need to extract word boundaries and syntactic rules in order to acquire the grammar of the language. In order to complete this task, the human brain analyzes the input at several levels (Uddén & Bahlmann, 2012). In the present study, we will focus on the level of syntactic analysis and the process of syntactic rule learning. This process has been hypothesized to depend on statistical learning, which might be based on abstract algebraic rules (Marcus, Vijayan, Rao, & Vishton, 1999) or predictive dependencies between words or phrases (Saffran, 2001, 2002).

### 5.1.1 The role of non-adjacent dependency learning in language acquisition

Syntactic structures that underlie language are hierarchical and therefore they cannot be learned purely based on associations between adjacent elements (Chomsky, 1957). In order to acquire the syntax of a language, one also needs to acquire the dependencies between non-adjacent words and morphemes (non-adjacent dependencies) which are abundantly present in natural language and often mark syntactic relations such as number and tense agreement (e.g. in English: *The dogs **are** bar**king*** and Dutch: *Wij **hebben** gisteren samen **gegeten** "Yesterday we **have** dine**d** together"*. Learning non-adjacent dependencies has proven to be a difficult task for adults (Bonatti, Pena, Nespor, & Mehler, 2005; Grama, Wijnen, & Kerkhoff, 2013; Newport & Aslin, 2004; Onnis, Monaghan, Richmond, & Chater, 2005; Peña, Bonatti, Nespor, & Mehler, 2002; Perruchet, Tyler, Galland, & Peereman, 2004; Van den Bos, Christiansen, & Misyak, 2012) as well as infants and children (Gómez, 2002; Gómez & Maye, 2005; Kerkhoff, De Bree, De Klerk, & Wijnen, 2013; Santelmann & Jusczyk, 1998; Van Heugten & Johnson, 2010). Both the linguistic level of analysis (Bonatti et al., 2005; Newport & Aslin, 2004) and the distance between the dependent elements (Grama et al., 2013) were shown to influence the degree to which participants are able to acquire the dependencies (for a review see Sandoval & Gómez, 2013). However, the nature of the computational mechanisms underlying the ability to acquire non-adjacent dependencies is still a topic of debate (Newport & Aslin, 2004; Peña et al., 2002; Perruchet et al., 2004; Van den Bos et al., 2012). In the present study, we ask which neural mechanisms underlie the acquisition of non-adjacent dependencies.

### 5.1.2 Artificial Grammar Learning in the brain

Behavioral studies have provided important insights with regard to the mechanisms involved in learning non-adjacent dependencies and the difficulties that learners face when acquiring these dependencies. A number of studies

have used Artificial Grammar Learning (AGL) paradigms to uncover the neural underpinnings of non-adjacent dependency learning (e.g. De Vries, Petersson, Geukes, Zwitserlood, & Christiansen, 2012; Uddén, Ingvar, Hagoort, & Petersson, 2012). Implicit Artificial Grammar Learning (Reber, 1967) has been used in many studies addressing language development and second language learning because AGL paradigms allow researchers to control a higher number of variables compared to a natural language learning setting.

Although AGL cannot represent the full complexity of learning a natural language, it is the paradigm of choice when studying neural correlates of syntactic development, because syntactic processing of previously learned artificial grammars has demonstrated a considerable overlap in neural activation with natural language processing in the posterior part of the left inferior frontal gyrus (BA 44/45), including Broca's area (Petersson, Folia, & Hagoort, 2012; Petersson & Hagoort, 2012). Previous AGL studies addressing non-adjacent dependency learning have mostly used complex artificial languages which participants learned over a longer period of time based on visually presented stimuli and have measured brain activation after the learning process was completed. This approach allows for consolidation, which has proved to be important in syntactic development (Nieuwenhuis, Folia, Forkstam, Jensen, & Petersson, 2013), but does not reveal how the brain reflects learning from mere exposure in real time. Although listening to a speech stream requires more sequential processing compared to visual presentation, the neural correlates of artificial grammar learning during auditory exposure to a speech stream have hardly been studied. In the present study, we use functional MRI in a highly controlled auditory artificial grammar learning paradigm to bridge this gap and to address on-line learning of non-adjacent dependencies.

The present experiment does not require participants to learn a complex artificial grammar with an extended vocabulary. However, computing transitional probabilities between adjacent elements is not sufficient for the acquisition of non-adjacent dependencies, which is expected to require complex unification operations at the phrasal level. Therefore, we hypothesize that acquiring non-adjacent dependencies from auditory exposure activates brain regions which are known to be involved in structural computations (for a review see Petersson & Hagoort, 2012). Based on an extensive body of research investigating the neural mechanisms underlying syntactic processing and artificial grammar learning (e.g. Bahlmann, Schubotz, & Friederici, 2008; Opitz & Friederici, 2003; Petersson et al., 2012; Uddén & Bahlmann, 2012; Uddén et al., 2008), we expect the Left Inferior Frontal Gyrus (LIFG) to be involved both during the exposure phase where the artificial grammar is processed for the first time and during grammatical judgment. In addition, we expect activation of the Superior Temporal Gyri (STG) during mere auditory exposure, as bilateral STG is hypothesized to be involved in speech perception (Hickok & Poeppel, 2000). Prolonged auditory exposure is expected to induce changes in

activation over time in both of these structures due to habituation in the auditory cortex and an increase in recognition of non-adjacent dependencies in LIFG.

### 5.1.3  Individual differences in brain and behavior

Earlier studies addressing neural correlates of second language acquisition (Tettamanti et al., 2002; Veroude, Norris, Shumskaya, Gullberg, & Indefrey, 2010) as well as birdsong learning (Bolhuis et al. (2000); Van der Kant et al. (2013), see also Chapter 3 of this thesis), have uncovered a relation between neural activation in response to the processing of vocalizations and individual differences in learning success. Based on these previous results, we will address the relation between individual differences in learning capacities and neural activation during learning through mere auditory exposure. Individual differences in learning ability are expected to correlate with neural activation especially in LIFG.

Working memory capacity has been suggested to influence syntactic learning both on the behavioral (Baddeley, 2003; Ellis, 1996) and on the neural level (e.g. Santi & Grodzinsky, 2007), and is thus expected to play a role in the detection and acquisition of non-adjacent dependencies. Furthermore, previous studies (Birdsong & Molis, 2001; Johnson & Newport, 1989) demonstrated age effects on adult second language acquisition, suggesting that age might play a role in the ability of participants in the present study to acquire non-adjacent dependencies. Potential correlations between learning outcome and neural activity might thus be influenced by individual variation in working memory capacity and age, which will be considered as factors in the present study.

## 5.2  Methods

### 5.2.1  Ethical statement

All participants gave written informed consent prior to inclusion in this study and were financially compensated for their participation. In accordance with Leiden University Medical Center (LUMC) policy all anatomical scans were reviewed by a radiologist. This screening yielded no anomalous findings. All experimental procedures were conducted under approval of the Medical Ethical Committee of the Leiden University Medical Centre, The Netherlands (CME no. NL42690.058.12).

### 5.2.2  Participants

Twenty healthy adult volunteers (12 males, mean age 28, range 18-43) were included in the present study. All participants were right-handed and reported normal hearing and normal or corrected-to-normal vision. They re-
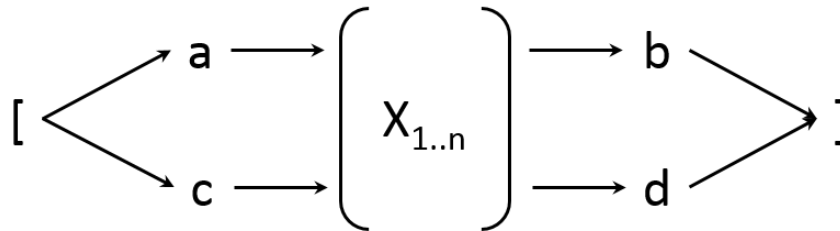
**Figure 5.1: Artificial grammar containing non-adjacent dependencies** Every X can occur in the middle position, but a is necessarily followed by b and c is necessarily followed by d. In the control condition, this dependency is absent.

ported no history of speech- or language disorders, other cognitive impairments or neurological damage. All participants were native Dutch speakers who used Dutch as their main language in daily communication and grew up in a monolingual Dutch environment.

### 5.2.3 Materials

For the present study, we employed an artificial grammar learning (AGL) paradigm where participants were first exposed to a learning set of the grammar described below and consequently tested on learning and generalization of the grammatical rules with a grammatical judgment task. The Artificial Grammar Learning paradigm enables us to study language learning in a more controlled way compared to natural language learning.

Participants were auditorily presented with phrases from an artificial grammar containing non-adjacent dependencies (NAD) and from a control language which had the same syllable structure and phonology, but lacked these dependencies. Stimuli were adapted from earlier behavioral studies on non-adjacent dependency learning by Dutch-speaking adults and children (Grama et al., 2013; Kerkhoff et al., 2013) and infant learners of English (Gómez, 2002) and were compatible with Dutch phonotactic constraints. Both the NAD and the control language consisted of strings of three pseudo-words. In the NAD language, a dependency exists between the first and last pseudo-word in each string, taking the form of aXb and cXd phrases (Figure 5.1). In the control language, there was no dependency between the first and last pseudo-word, resulting in strings where any of the a, b, c, and d-elements could occur both in the initial and in the final position and could be combined with any other of the a, b, c, and d-elements, maintaining the syllable structure of the strings. To allow for assessment of learning of the NAD language, ungrammatical stimuli containing a violation of the dependency were constructed. In these stimuli, the dependency between the first and last syllable was switched, resulting in aXd and cXb phrases.

Two sets of monosyllabic a, b, c and d-elements were used, where for each

|          | **NAD stimuli** | **Control stimuli** | **Ungrammatical stimuli** |
|----------|-----------------|---------------------|---------------------------|
| Group 1  | mo noeba zop<br>tep noeba ra<br>mo wadim zop<br>tep wadim ra | poe noeba jik<br>si noeba poe<br>jik wadim si<br>lut wadim jik | mo noeba ra<br>tep noeba zop<br>mo wadim ra<br>tep wadim zop |
| Group 2  | poe noeba jik<br>lut noeba si<br>poe wadim jik<br>lut wadim si | mo noeba zop<br>ra noeba zop<br>zop wadim tep<br>tep wadim ra | poe noeba si<br>lut noeba jik<br>poe wadim si<br>lut wadim jik |

**Table 5.1: Examples of NAD and control stimuli for auditory presentation** For illustration purposes, only a subset of the X elements occurring in the full stimulus set is used. The two groups only differentiate in the stimuli they hear and are not analyzed separately. NAD: artificial grammar containing non-adjacent dependencies, Control: artificial grammar lacking non-adjacent dependencies.

participant one set was assigned to each language (NAD or Control) to correct for any effects of acoustic differences between the sets. A set of in total 119 bisyllabic X-elements was constructed and combined with the a, b, c and d-elements to construct the full set of stimuli. A latin-square design was employed to counterbalance assignment of the stimuli to participants. Stimuli containing each of the sets of abcd-elements were either assigned to the NAD or to the control language for each participant. Moreover, X-elements were assigned to either the exposure or the test phase for each participant. A number of examples of stimuli from the two languages and of violations are given in Table 5.1.

According to earlier findings (Kerkhoff, personal communication; Peña et al. (2002)), rule learning only occurred the pseudo-words within a phrase were separated by short pauses. Therefore, we inserted 250 ms pauses between the pseudo-words in each stimulus. All pseudo-words were pronounced in carrier sentences by a Dutch-speaking male adult and recorded using Adobe Audition (Adobe Systems, San Jose, California) and a Sennheiser directional microphone (Sennheiser electronic, Wedemark, Germany). They were then isolated from the speech stream and concatenated into the final stimuli using Praat software (Boersma, 2002), version 5.3.41, freely available from http://www.praat.org/. The final stimulus inventory contained 476 NAD stimuli, 476 control stimuli and 476 ungrammatical stimuli. Although more combinations were possible for the control condition, a subset was selected for the stimulus inventory that balanced frequency and position between a, b, c and d-elements.

### 5.2.4 Experimental paradigm

In this two-phase fMRI experiment, participants learned an artificial grammar containing non-adjacent dependencies from mere auditory exposure. During an initial exposure phase, participants listened to the artificial grammar (NAD) and the control language, which were presented through pneumatic headphones, which were part of the scanner intercom system (Philips Healthcare, Best, The Netherlands). In a grammatical judgment task, which followed the exposure phase, participants listened to items that either followed or violated the artificial grammar they had been exposed to. Participants were then asked to judge the grammaticality of these items, thus testing learning and generalization of the artificial grammar. Stimulus presentation was controlled and responses were recorded using E-prime (Psychology Software Tools, Sharpsburg, Pennsylvania).
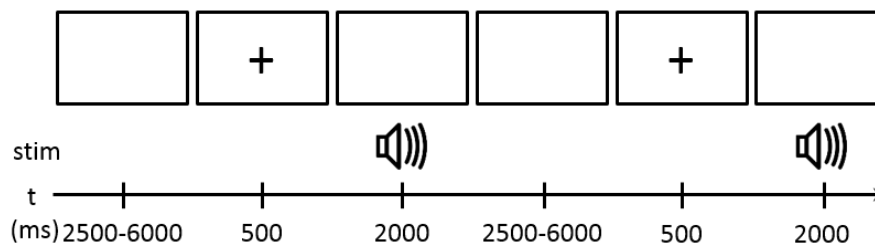


**Figure 5.2: Experimental paradigm for the exposure phase** Auditory stimuli ('stim') are played through headphones during presentation of a black screen for 2 s.

During the exposure phase, participants listened to the two artificial grammars (NAD and control) which were auditorily presented to them. Participants were presented with a total of 100 stimuli during two blocks of approximately six minutes, with a self-paced pause between the blocks. A different subset from the inventory of NAD stimuli was presented to each participant and stimuli were not repeated. Stimulus presentation followed an event-related paradigm with a stimulus duration of 2,000 ms and a variable inter-stimulus interval of 2,500-6,000 ms and presentation of each stimulus was preceded by a fixation cross presented on the screen for 500 ms. Behavioral pilots showed that random presentation of items from the two languages did not result in any learning (above-chance performance on grammaticality judgments). Therefore, stimuli from the NAD and control languages were alternated in groups of six stimuli, within which selection of items from the language was pseudo-random in the sense that stimuli could not be presented more than once per participant. During both blocks of the exposure phase a GE-EPI time-series was acquired.

Following the exposure phase, participants were asked to judge phrases from the NAD language and violations of the NAD language on their grammaticality in a test phase. A total of 66 stimuli were presented through head-

phones in two blocks of approximately five minutes with a pause between blocks. Participants were asked to indicate for each stimulus whether the phrase belonged to the language they had heard during the exposure phase. Participants were given a 3 s time window, which was indicated with a question mark on the screen, to respond with a button press of the right or left index finger. Stimuli were presented in an event-related paradigm using the same timing parameters as in the exposure phase. In the grammatical judgment task, stimulus presentation was randomized with no grouping of NAD or ungrammatical stimuli. Participants were never presented with the same NAD phrases in the exposure phase and test phase. To further assess generalization of the grammar, part of the stimuli in the test phase contained X-items that the participants had not heard in the exposure phase. Selection of the NAD phrases and X-items that were only presented in the test phase was counterbalanced between participants. During both blocks of the test phase, fMRI data were acquired.

In addition to the fMRI experiment, participants were tested on their working memory capacity using both a verbal and a non-verbal working memory task (WAIS-IV forward and backward digit span (Pearson, The Netherlands) and Exef mental counters (Department of Developmental Psychology, University of Amsterdam, The Netherlands)). All participants completed a questionnaire on their language background, followed by a questionnaire assessing the strategies participants had used and the regularities they had detected during the AGL task.

### 5.2.5  MRI data collection and analysis

All MRI data were acquired at the LUMC using a 3 Tesla Philips Achieva TX MRI system (Philips Healthcare, Best, The Netherlands) with a whole-head receiver coil. Participants were screened for counter-indications before entering the scanning room. In order to minimize movement and increase the comfort level of the participant, cushioning was placed around the head and ear plugs and headphones were provided for hearing protection.

Both during the exposure phase and the test phase, two fMRI time series of 130 $T_2^*$-weighted whole-brain single-shot GE-EPI volumes were acquired for each participant, including two dummy scans to allow the longitudinal magnetization to reach equilibrium (35 transverse slices, TR/TE: 2550/30 ms, flip angle: 80°, voxel size 2.75 mm$^3$, including a 10% inter-slice gap, field of view (FOV): 220 x 217.25 x 96.25 mm). The slightly longer TR and the smaller number of slices compared to the EPI sequence used for the resting state scans were selected in order to minimize the gradient noise level. To further reduce the noise level inside the scanner during presentation of the auditory stimuli, the gradient slew rate was increased using the Philips soft-tone parameter (factor 5), which has shown to reduce sound pressure levels with 12dB (Rondinoni, Amaro Jr, Cendes, Dos Santos, & Salmon, 2013).

Preprocessing of the functional data acquired during both the exposure

phase and the grammatical judgment task was carried out using SPM8 (Wellcome Trust Centre for Neuroimaging, London, UK; http://www.fil.ion.ucl.ac.uk/spm/). To allow for accurate localization of the functional activations, the high-resolution $T_1$-weighted anatomical scan of each subject was segmented and registered to the Montreal Neurological Institute (MNI) $T_1$-template. Each fMRI time series was realigned to correct for head movement and it was confirmed that head movements did not exceed 5 mm in the x, y or z direction. Functional images were then co-registered to the high-resolution anatomical scan of the same participant and normalized to the MNI template using the parameters obtained from the registration of the high-resolution anatomical scan. Finally, the fMRI images were smoothed with an 8 mm full-width-at-half-maximum (FWHM) isotropic Gaussian kernel.

Statistical analyses of the fMRI data at the individual and group level were performed using the General Linear Model implemented in SPM8. Data were modeled as a series of events and convolved with a canonical hemodynamic response function (HRF) and low-pass filtered at 128 Hz. For both the exposure phase and the grammatical judgment task, the start of presentation of each auditory stimulus was modeled as an event of interest. Separate regressors were defined for NAD and control trials from each session in the exposure phase. For the grammatical judgment task, separate regressors were defined for correct and incorrect trials and NAD and ungrammatical trials, respectively, resulting in a total of four regressors. Regressors were entered into a group level random-effects analysis as separate contrast images for each individual. Statistical differences between stimulus-evoked BOLD signals were assessed using a one-way repeated measures ANOVA followed by post-hoc one-tailed t-tests. Regression analyses with behavioral results from the test phase and behavioral post-tests were also performed to assess potential brain-behavior correlations. For all group analyses $p$-values were corrected for multiple tests using the Family Wise Error method based on the Random Field Theory (Worsley et al., 1996) on the cluster level and a cluster threshold of five voxels was applied.

## 5.3 Results

### 5.3.1 Behavioral results

Data from the grammatical judgment task which was administered in the scanner after exposure showed that most participants had major difficulties learning the non-adjacent dependencies in the NAD grammar. Participants performed significantly above chance on the rejection of violations (one-tailed: $T = 1.79$, $p = 0.045$), while performance on accepting NAD phrases was below chance level (Table 5.2). A paired t-test showed that accuracy scores differed significantly between NAD phrases and violations (two-tailed: $T = -3.1$, $p = 0.005$). However, no significant difference was found between reaction times

to NAD phrases and violations (paired t-test: $T$ = -0.41, $p$ = 0.69, see Table 5.2 for means).

|               | **% Correct responses** Mean (sd) | **Response time (ms)** Mean (sd) |
|---------------|-----------------------------------|----------------------------------|
| NAD phrases   | 39 (18)                           | 955 (310)                        |
| Violations    | 58 (19)                           | 968 (337)                        |
| Total         | 48 (13)                           | 961 (317)                        |

**Table 5.2: Behavioral performance on the grammatical judgment task** Percentage of correct responses is given. For NAD stimuli, a response is correct when the participant accepts the phrase, while rejecting the phrase is the correct response for violations. 50% correct represents chance level.

### 5.3.2   Exposure phase

Unless stated otherwise, all fMRI results are FWE-corrected for multiple comparisons at the cluster level with corrected $p$-values below $p$ = 0.05 considered significant. $R^2$-values are computed based on the highest T-value within the reported cluster from the SPM correlation map. MNI coördinates, $Z$-values and cluster extents for the results reported in this section can be found in Tables 5.3 and 5.4.

F-tests on the group level did not reveal a main effect of session. With a threshold of $p$ < 0.001, a small cluster in the left Insula showed a session*stimulus interaction, but this did not survive correction for multiple comparisons ($Fmax$ = 14.06, $p$-FWE = 0.999, $p$-uncorr < 0.001). Because no differences between sessions were found, we considered both sessions together in a paired t-test for stimulus effects. Both NAD minus Rest and control minus Rest showed activated clusters in bilateral STG (NAD > Rest: LSTG $p$ < 0.001; RSTG $p$ < 0.001; control-Rest: LSTG $p$ < 0.001; RSTG $p$ < 0.001) and IFG (NAD > Rest: LIFG $p$ < 0.001; RIFG $p$ < 0.001; Control > Rest: LIFG $p$ = 0.002; RIFG $p$ < 0.001). F-tests on the group level did not reveal a significant main effect of stimulus (NAD vs. control) for the exposure phase. Clusters shown in the right posterior and middle cingulate gyrus at a threshold of $p$ <0.001 did not survive FWE-correction (Post. Cing.: $p$-FWE = 0.997, $p$-uncorr < 0.001; Mid. Cing.: $p$-FWE = 0.96, $p$-uncorr < 0.001). See Figure 5.3 for activated clusters and Table 5.3 for peak $Z$-values and cluster extents.

In line with previous fMRI results in humans (e.g. McNealy, Mazziotta, & Dapretto, 2006; Veroude et al., 2010) and songbirds (see chapter 3 and 4 and Van der Kant et al., 2013), we employed regression analysis to investigate individual differences in brain and behavior. Performance on rejection of violations was used as an indicator of sensitivity to the non-adjacent dependencies in the grammar, because participants scored below chance level on
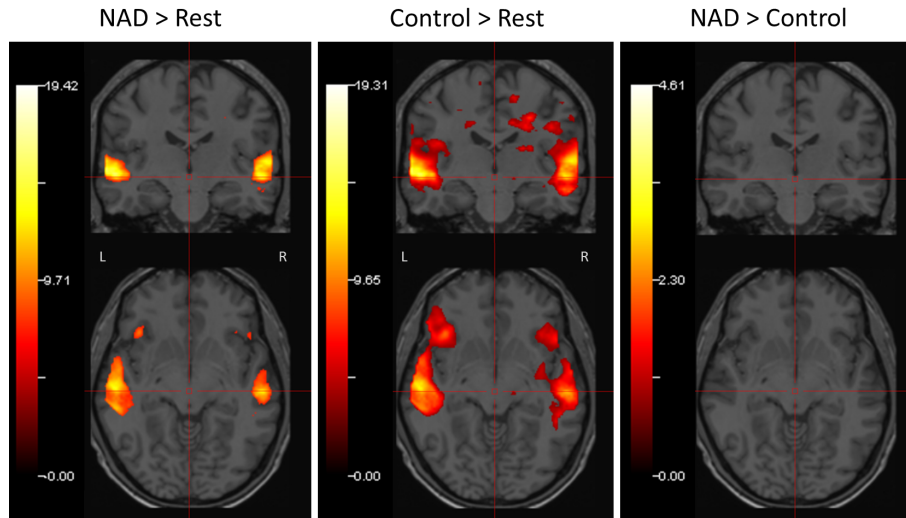
**Figure 5.3: Brain activation during exposure to the NAD and Control languages, FWE corrected for the whole brain with a $p < 0.05$ 5-voxel threshold.** Left: Activation in response to the language containing non-adjacent dependencies compared to Rest. Middle: Activation in response to the language without non-adjacent dependencies compared to Rest. Right: Differential activation induced by NAD compared to Control does not survive whole-brain FWE correction.

the acceptance of NAD phrases. Regression analyses showed a positive correlation between the success with which individual subjects were able to reject violations of the grammar (% correct) after exposure to the NAD and control grammars and neural activation in response to the NAD grammar compared to the control condition in the Left Inferior Frontal Gyrus (LIFG) ($R^2 max = 0.69$, $p = 0.004$) and the bilateral Superior Temporal Gyri (STG) and Insulae (In) (Left: $R^2 max = 0.79$, $p < 0.001$; Right: $R^2 max = 0.65$, $p < 0.001$). A trend towards such a correlation was found in RIFG ($R^2 max = 0.7$, $p = 0.09$). See Figure 5.4 for clusters and correlation plots and Table 5.4 for $R^2$, $Z$-values and cluster extents. Correlations with NAD minus Rest were neither significant in IFG (LIFG: $R^2 max = 0.73$, $p = 0.206$; RIFG: $R^2 max = 0.48$, $p = 0.998$) nor in STG (LSTG: no clusters, RSTG: $R^2 max = 0.53$; $p = 0.992$) and In (L In: $R^2 max = 0.49$, $p = 0.771$; R In: $R^2 max = 0.53$, $p = 0.121$). No correlations with NAD > Control were found for age or any of the behavioral variables digit span and mental counters.

### 5.3.3 fMRI results grammaticality judgment task

MNI coördinates, $Z$-values and cluster extents for the results reported in this section can be found in Table 5.5. Potential differences between the two sessions of the grammatical judgment task were first explored using a random-effects analysis. Because no significant effect of session (strongest cluster in
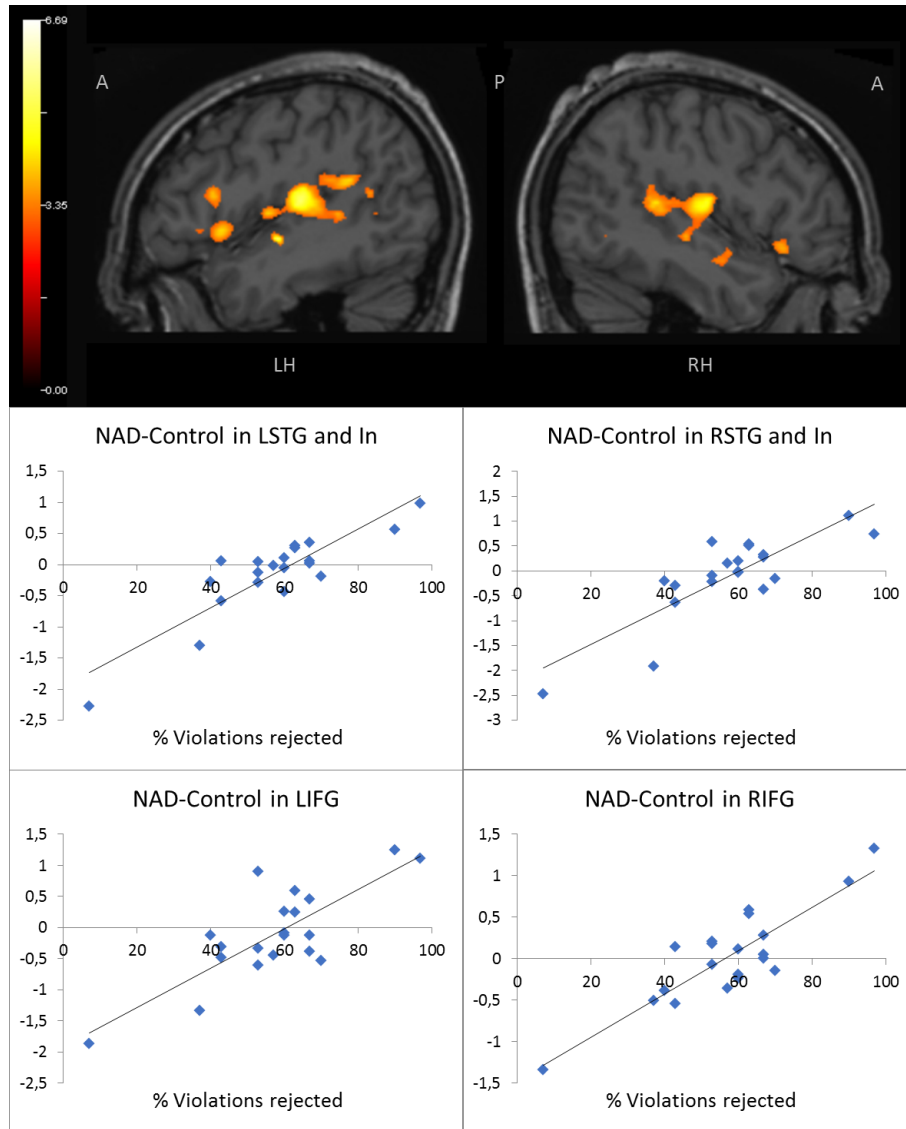
**Figure 5.4: Correlation of NAD > Control activation with the percentage of correctly rejected violations in the grammatical judgment task.** Top: Bilateral STG / Ins and LIFG show a correlation between the differential activation in response to NAD stimuli during the exposure phase and the number of correctly rejected violations. Bottom: Plots illustrating the correlation between the total NAD > Control activation within each cluster shown on the top and the performance on violation rejection in the grammatical judgment task. These plots are shown for illustration purposes and are not used as a source for the $R^2$-values reported in this section.

| Contrast | Brain region | L/R | MNI coördinates | | | $Z$ | Vox. |
|---|---|---|---|---|---|---|---|
| | | | x | y | z | | |
| NAD | STG / In. | L | -54 | -22 | -4 | 6.96 | 1667 |
| minus | | R | 66 | -16 | -2 | 7.52 | 2173 |
| Rest | IFG | L | -44 | 24 | -12 | 5.35 | 41 |
| | | R | 40 | 26 | -8 | 5.89 | 151 |
| | Front. Par. | R | 40 | 22 | 22 | 5.82 | 135 |
| Control | STG / In. | L | -56 | -22 | -6 | 6.64 | 1442 |
| minus | | R | 66 | -16 | -2 | 7.51 | 1986 |
| Rest | IFG | L | -44 | 22 | -12 | 5.10 | 7 |
| | | R | 44 | 22 | -18 | 5.64 | 141 |
| | Ant. Cing. | R | 24 | 24 | 20 | 5.53 | 28 |
| | Post. Cing. | R | 28 | -30 | 34 | 6.17 | 54 |

**Table 5.3: Brain activation during exposure to the NAD and Control languages.** Cluster extents are based on the FWE-corrected T-map with a $p < 0.05$ 5-voxel threshold as shown in Figure 5.3. L/R = Left/Right hemisphere, Vox. = cluster extent in voxels, $Z$ = $Z$-value of the peak voxel in the specified brain region, NAD = non-adjacent dependencies language, Control = language without non-adjacent dependencies, Stim = stimulus, Ant. / Post. Cing. = Anterior / Posterior Cingulate cortex, STG = Superior Temporal Gyrus, In. = Insula, IFG = Inferior Frontal Gyrus.

Right posterior Cingulate gyrus: $p = 0.324$) or session*stimulus interaction was present within the test phase, the two sessions were considered together in further analyses. One-sample t-tests for NAD phrases minus Rest and violations minus Rest showed significant activation for violations but not NAD phrases (violations Anterior Insula: peak-level $p = 0.001$; NAD: no clusters). $F$-tests for differences between NAD phrases and violations did not show significant activations, two clusters in the bilateral Middle Temporal Gyri did not survive correction for multiple comparisons (Left: $p$-FWE = 0.924, $p$-uncorr < 0.001; Right: $p$-FWE = 0.489, $p$-uncorr < 0.001). However, a trend towards a main effect of accuracy was shown in the Left inferior frontal gyrus (peak-level $p = 0.066$).

To further investigate the origin of this trend a post-hoc t-tests was performed, showing a differential activation for Correct minus Incorrect responses in LIFG (peak-level $p = 0.033$). The difference in activation between correct and incorrect responses was significant in LIFG for violations of the NAD language ($p = 0.003$; temporal pole cluster did not reach significance: peak-level $p = 0.099$), but was absent for NAD phrases containing grammatical non-adjacent dependencies (no clusters). However, the stimulus*accuracy interaction in LIFG did not survive corrections ($p$-FWE = 0.73, $p$-uncorr < 0.001). It should be noted here that the only brain activity in response to the presentation of the stimulus and not to the response itself was measured and that no feedback on the response was given. Although participants did not learn the

| L/R | Voxels | Brain regions | MNI coördinates | | | $R^2$ | $Z$ |
|-----|--------|---------------|------|------|------|-------|------|
|     |        |               | x    | y    | z    |       |      |
| L   | 2022   | STG / Insula  | -48  | -22  | 12   | 0.79  | 4.71 |
|     |        | Angular Gyrus | -46  | -44  | 20   | 0.59  | 3.93 |
|     |        | IFG           | -44  | 22   | -2   | 0.69  | 4.54 |
| R   | 1150   | STG / Insula  | 40   | -14  | 10   | 0.65  | 4.67 |
|     |        | Angular Gyrus | 58   | -30  | 18   | 0.58  | 3.92 |
|     | 128    | IFG           | 42   | 26   | -6   | 0.70  | 4.62 |

**Table 5.4: Correlation between the success in rejecting violations and NAD > Control differential activation during the exposure phase** Cluster extents and $R^2$ values are based on the uncorrected T-map with a $p < 0.0001$ 5-voxel threshold as shown in Figure 5.4. L/R = Left/Right hemisphere, Voxels = cluster extent in voxels, STG = Superior Temporal Gyrus, IFG = Inferior Frontal Gyrus.

language in the behavioral sense, the brain was sensitive to violations.
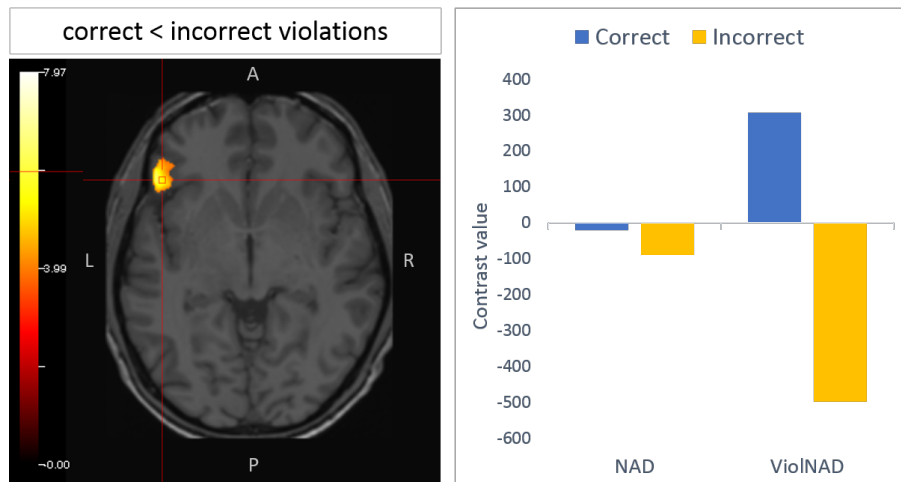


**Figure 5.5: Activation and contrast values shown during the grammatical judgment task.** Left: LIFG shows differential activation for (correctly) rejected compared to (incorrectly) accepted violations. Right: Contrast values for correct and incorrect trials per stimulus type.

## 5.4 Discussion

The present study aimed to unravel how the brain reflects online learning of an artificial grammar containing non-adjacent dependencies. The behavioral results showed that overall, participants scored significantly above chance

| Contrast | Brain region | L/R | MNI coördinates | | | Z | Vox. |
|---|---|---|---|---|---|---|---|
| | | | x | y | z | | |
| ME session | Post. Cing. | R | 20 | -38 | 32 | 4.13 | 16 |
| viol. - Rest | Ant. Insula | L | -28 | 28 | 12 | 5.60 | 146 |
| NAD vs viol. | Mid. Temp. | L | -48 | -42 | -8 | 3.66 | 34 |
| | | R | 50 | -30 | -12 | 4.06 | 75 |
| ME accuracy | IFG | L | -50 | 36 | -12 | 4.66 | 93 |
| Correct-Incorrect | IFG | L | -50 | 36 | -12 | 4.80 | 133 |
| Corr-Incorr viol. | IFG | L | -50 | 36 | -12 | 5.92 | 308 |
| | Temp. pole | L | -54 | 4 | -26 | 4.53 | 86 |
| Stim*Accuracy | IFG | L | -50 | 36 | -12 | 3.87 | 40 |

**Table 5.5: Brain activation during the grammatical judgment task** Values are uncorrected with a $p < 0.001$ 5-voxel threshold. L/R = Left/Right hemisphere, Vox. = cluster extent in voxels, ME = main effect, viol. = violations, NAD = NAD phrases, Stim = stimulus, Post. Cing. = posterior cingulate cortex, Ant. = Anterior, Mid. = Middle, Temp = Temporal, IFG = Inferior Frontal Gyrus.

when rejecting violations but had difficulties judging phrases as grammatical that belonged to the grammar that they were previously exposed to. Because earlier behavioral pilots yielded higher grammatical judgment scores than the grammatical judgment task in the scanner, we can assume that despite our efforts to reduce scanner noise, learning the grammar was more challenging in the noisy scanner environment.

In the grammatical judgment task, we observed considerable individual differences both in the ability to accept grammatical NAD phrases and to reject violations. These individual differences were reflected in a correlation between the success with which participants rejected violations during the grammatical judgment task and the differential activation induced by the NAD language compared to the control language during the exposure phase in the bilateral STG and Insulae and the left Inferior frontal gyrus. Furthermore, during the grammatical judgment task, correct responses induced more activation in LIFG than incorrect responses for violations but not for NAD phrases.

### 5.4.1 Differential activation as a precursor for learning

The correlation between brain activation and learning success found in our data shows that the LIFG and the bilateral STG and Insulae are involved in implicit learning of non-adjacent dependencies through auditory exposure. Involvement of LIFG has been previously demonstrated in numerous artificial grammar learning studies (e.g. Folia & Petersson, 2014; Uddén & Bahlmann, 2012; Yang & Li, 2012), which among other also addressed hierarchical dependencies (Bahlmann et al., 2008). However, the role of (particularly the right) STG and Insula appears to be less prominent in most artificial grammar learn-

ing studies. This could either be a result of the modality of the present experiment, where all stimuli were presented auditorily, or of the fact that data were collected during as opposed to after exposure to the artificial grammar. Activation of the auditory cortex during the presentation of auditory stimuli is not surprising, but it should be noted that the bilateral STG is differentially activated by the NAD compared to the control language. Participants were given no clues that they were listening to two different languages and NAD and control stimuli were only distinguishable based on the higher auditory variation provided by the control language, which occurred because different combinations between first and last elements were possible. This variation could certainly be reflected in the differential activation induced by NAD compared to control stimuli. However, the amount of variation did not differ between participants while the differential activation showed considerable individual variation. Moreover, like in many AGL studies where fMRI data were acquired after learning, only LIFG activation was shown during the grammatical judgment task. Therefore, it is likely that both modality and data collection during learning play a role.

In addition to the involvement of LIFG, the Insulae and bilateral STG in artificial grammar learning, the correlation between behavioral learning measures and differential brain activation also indicates that individual differences in sensitivity to violations are already reflected in the brain during the learning process. Because the correlation with behavior in IFG, STG and In is only present for the differential (NAD > Control) activation, it is unlikely that these correlations are due to individual differences in auditory capacity or attention. Although attention and active processing has been suggested to play a large role in the ability to learn non-adjacent dependencies (Pacton & Perruchet, 2008), attentional factors cannot explain our results, because the NAD and control languages were not discernable during exposure and were thus unlikely to show differences in terms of attention. Because we were able to collect fMRI data online during learning and participants provided us with grammatical judgments after exposure, it is likely that the increased differential activation in more successful learners is a precursor rather than a result of learning. This interpretation is further supported by the fact that, although no significant difference between sessions was shown, most differential activity for the NAD compared to the control language in more successful learners was seen in the first session. Moreover, the individual differences in activation in an experiment where exposure times are short indicate that individual differences in language learning might result from differences in neural recruitment that develop early in the learning process.

The behavioral data also showed that after the limited amount of exposure to the NAD language, the learning process was not completed for most participants. The majority scored at or just above chance on the grammatical judgment task. Furthermore, in the post-test questionnaire administered after the grammatical judgment task, participants did not report any awareness of

the non-adjacent dependencies present in the NAD grammar. It is therefore highly unlikely that the increased differential activation in participants with higher scores in the grammatical judgment task was due to the use of explicit learning strategies by more successful learners. In accordance with previous studies (McNealy et al., 2006; McNealy, Mazziotta, & Dapretto, 2010), these results show neural reflections of artificial grammar learning before learning can be reliably determined using behavioral methods. Similar correlations between learning success and neural activation or connectivity have been found in language learning studies (Veroude et al., 2010; White, Genesee, & Steinhauer, 2012) and in studies of birdsong learning (see Bolhuis et al. (2000); Van der Kant et al. (2013) and Chapter 4 of this thesis). These results suggest that increased neural activation during sensory stimulation might aid learning across modalities and species.

### 5.4.2   Sensitivity to the deviant as an early stage in learning

From the behavioral as well as the fMRI results from the grammatical judgment task it was apparent that violations of the NAD language served as a more salient stimulus compared to grammatical NAD phrases. In the grammatical judgment task, violations showed a higher response accuracy. Moreover, the differential activation in LIFG for correctly compared to incorrectly judged phrases was present for violations, but not for NAD stimuli. These results show that violations are processed differently from NAD phrases, however, because the segmental content of the languages is counterbalanced between participants, this difference in processing is purely a result of the knowledge about the NAD language obtained during the exposure phase. These findings are in accordance with earlier results which show that violations in a learned artificial grammar activate LIFG (Petersson, Forkstam, & Ingvar, 2004).

   Both the stronger neural activation and the higher performance on the violations might result from the more general neural mechanism of deviance detection. Because participants have been exposed to 100 phrases containing the same non-adjacent dependencies, expectancies about similar stimuli have formed in the brain, which are strong enough to induce an error signal when violated. In contrast, the rules have not been learned to such an extent that a grammatical phrase can be reliably classified as such. The mechanism of deviance detection is used in non-linguistic (MMN), semantic (N400) as well as syntactic (P600) ERP studies to investigate the detection of violations. The amplitude of the P600 during grammatical judgment in second language acquisition has been shown to correlate with behavioral performance on the grammatical judgments (White et al., 2012), further illustrating the role of error signals in language learning. Deviance detection is also used in behavioral studies of infant language learning, where babies are expected to show more interest in new or deviant stimuli after exposure to a string of similar stimuli (several paradigms are discussed in Aslin & Fiser, 2005). This might suggest

that the increased sensitivity to violations might be the result of a deviance detection mechanism, which aids language learners in very early stages of learning.

The differential activation in LIFG induced by correctly rejected compared to incorrectly accepted violations in the grammatical judgment task shows an increased sensitivity to ungrammatical compared to grammatical NAD phrases, for which no such difference was found. Correctly rejected violations induced activation of the LIFG and accepted violations induced de-activation of the same area, while listening to grammatical NAD phrases induced minor BOLD signal changes in LIFG. Although we should be cautious interpreting the de-activation of LIFG in response to violations that were consequently not rejected, the activation of LIFG in response to correctly rejected violations might represent the detection of a deviant dependency structure based on the knowledge of the NAD language that has been acquired during the exposure phase. This interpretation is in line with the finding of Hein et al. (2007) that in an audiovisual integration task incongruent stimuli activate IFG. The neural dissociation found between NAD and control stimuli in the exposure phase and the neural sensitivity to violations during the test phase without awareness of any differences suggests that these sensitivities developed based on statistical computations (Turk-Browne, Scholl, Chun, & Johnson, 2009).

### 5.4.3 Conclusions

The present study uncovered a number of interesting connections between brain and behavior during artificial grammar learning, which show how individual differences in artificial grammar learning and possibly natural language learning are reflected in the brain.

First, we showed that the extent to which participants were able to deduce non-adjacent dependencies from an artificial grammar correlated with the strength of activation in brain areas that are implicated in artificial grammar learning and natural language processing. This result suggests that neural activation in the language network during learning might predict artificial grammar learning and possibly also natural language learning outcomes.

Participants were found to be more sensitive to violations in the NAD language and showed a differential activation for rejected > accepted violations but not NAD phrases during the grammatical judgment task. Both these findings and the brain-behavior correlation were in absence of reliable behavioral learning. This suggests that the brain reflects early stages of learning, which might include the development of sensitivity to violations of regularities that are present in the input.

CHAPTER 6

---

# Reflections of individual differences in grammar learning on the language network

---

*How do individual differences in structural and functional connectivity in the human language network influence adult language learning? Both structural and functional connectivity have been reported to show deficiencies in patients with cognitive or language disabilities. However, few studies to date have studied how individual differences in brain connectivity relate to language learning abilities in healthy participants. The present study employed a combination on functional connectivity analysis of resting-state fMRI (rs-fMRI) and Diffusion Tensor Imaging (DTI) in order to investigate the relation between individual differences in brain connectivity and artificial grammar learning ability. Results showed that both structural and functional connectivity between left hemisphere IFG Pars Triangularis, part of Broca's area, with posterior regions in the language network correlated with the ability of individual participants to acquire an artificial grammar. These results suggest that individual differences in brain connectivity, especially between Broca's area and other regions important for language processing, underlie individual differences in language learning abilities.*

## 6.1   Introduction

Although we share our ability for vocal learning with other species, humans are unique in their ability to spontaneously learn language, which exhibits a complex syntactic system. Learning the syntactic system of our native language takes up a large part of our development and is a fragile process, which

can be impacted by developmental disorders or brain trauma. Human infants are equipped to acquire a highly complex syntactic system without instruction, whereas adult learners have more difficulties acquiring a new grammatical system and show considerable individual differences in the ability to acquire new syntactic structures.

How do individual differences in the human ability to learn a new grammar relate to the neural systems subserving language processing? In a previous fMRI study (see Chapter 5), including the same group of participants that was included in the present study, participants listened to phrases from an artificial grammar containing non-adjacent dependencies, followed by a grammatical judgement task. FMRI results showed a correlation between activation in the Left Inferior Frontal Gyrus (LIFG), the bilateral Superior Temporal Gyri and the bilateral Insulae and the ability to reject violations of the grammar in the grammatical judgement task. A similar effect was found in one of our previous studies assessing the perception of learned vocalizations in zebra finches, a songbird species (see Chapter 3. Individual differences in song learning were found to be reflected in activation of the auditory midbrain in response to auditory exposure to the learned vocalizations.

In humans, A growing body of literature shows that LIFG is an important brain region for syntactic processing in both natural language processing and artificial grammar learning (e.g. Folia & Petersson, 2014; Petersson et al., 2012; Petersson & Hagoort, 2012; Uddén & Bahlmann, 2012; Yang & Li, 2012). Our previous results suggest that activity in a large part of the network involved in language processing underlies individual differences in artificial grammar learning and possibly natural language learning abilities.

However, the brain operates in large-scale interconnected networks rather than individual regions. Moreover, the possibility that individual differences in the ability to learn this specific artificial grammar, containing dependencies between non-adjacent elements, were purely task-related, cannot be excluded based on the fMRI data. The neural basis of individual differences of artificial grammar learning should thus be investigated on the network level, rather than in isolated brain regions. In order to assess whether artificial grammar learning is related to individual differences in connectivity within the language network (Lohmann et al., 2010), we employed resting-state functional MRI (rs-fMRI) and Diffusion Tensor Imaging (DTI) to assess functional and structural connectivity between LIFG, including Broca's area and temporal and parietal language processing regions, including Wernicke's area. Functional connectivity changes induced by auditory exposure to the artificial grammar will be assessed by acquiring rs-fMRI data both before and after participants listen to the artificial language. Resting-state functional connectivity has been shown to reflect structural connectivity (Greicius, Supekar, Menon, & Dougherty, 2009). However, structural connectivity alone does not differentiate between inhibitory and excitatory connections, while functional connectivity between two regions does not necessarily mean there is a direct white

matter connection between the two. By combining DTI and rs-fMRI measurements, we aim to arrive at a detailed analysis of how connectivity within the language network is related to artificial grammar learning.

Prior studies have uncovered links between functional and structural connectivity and language impairments as well as individual differences in language functions of healthy participants. Individual differences in reading (Koyama et al., 2011) as well as learning words from an unfamiliar language (Veroude et al., 2010) have been associated with resting-state functional connectivity differences, while white-matter integrity was shown to be related to artificial grammar learning success (Flöel, De Vries, Scholz, Breitenstein, & Johansen-Berg, 2009). In patients diagnosed with primary progressive aphasia, impairment of syntactic comprehension was shown to correlate with the amount of damage to the left hemisphere dorsal language pathway (Wilson et al., 2011), while autistic individuals with language impairments exhibit decreased functional connectivity (e.g. Kana, Keller, Cherkassky, Minshew, & Just, 2006). We aim to show how functional and structural connectivity together might subserve artificial grammar learning in healthy adults.

The interconnected network subserving language processing and speech production in the left hemisphere shows different pathways (Saur et al., 2008), which subserve different language processing functions. A ventral pathway connects LIFG with anterior temporal regions through the Extreme Capsule, while a dorsal pathway connects LIFG to parietal and posterior temporal regions through the Superior Longitudinal and Arcuate Fasciculi. Prior rs-MRI data have shown that functional connectivity of language-related regions along these pathways is highly reproducible across participants (Tomasi & Volkow, 2012). Furthermore, several studies (for a review, see Friederici & Gierhan, 2013) have shown that the dorsal language pathway subserves complex syntactic operations. Therefore, individual differences in artificial grammar learning are mainly expected to be related to differences in structural connectivity in the Arcuate or Superior Longitudinal Fasciculus. Functional connectivity between regions connected by the dorsal pathway is also expected to reflect individual differences in artificial grammar learning. Additionally, functional connectivity within the language network is expected to show a larger increase between sessions acquired before and after exposure to an artificial grammar in participants who were more successful at learning the grammar.

## 6.2  Methods

### 6.2.1  Ethical statement

All participants gave written informed consent prior to inclusion in this study and were financially compensated for their participation. In accordance with Leiden University Medical Center (LUMC) policy, all anatomical scans were

reviewed by a radiologist. No anomalous findings were reported. All experimental procedures were conducted under approval of the Medical Ethical Committee of the Leiden University Medical Centre, The Netherlands (CME no. NL42690.058.12).

### 6.2.2  Participants

For the present study, the same participants were included that participated in the study described in Chapter 5. This group consisted of twenty healthy adult volunteers (12 males, 8 females, mean age 28, range 18-43). All participants were right-handed and reported no history of speech- or language disorders, other cognitive impairments or neurological damage. All participants were native speakers of Dutch.

### 6.2.3  Imaging Procedure

All imaging data were acquired at the LUMC using a 3 Tesla Philips Achieva TX MRI system (Philips Healthcare, Best, The Netherlands) with a whole-head receiver coil. Participants completed a safety questionnaire to screen for counter-indications before scanning. In order to minimize movement, cushioning was placed around the head and for the DTI scan participants were informed about possible vibrations in the scanner bed prior to scanning. Ear plugs and headphones were provided for hearing protection.

Resting-state (rs) fMRI time-series were acquired prior to and after auditory exposure to an artificial language containing non-adjacent dependencies and a control language in order to determine how artificial grammar learning influences functional connectivity between brain areas known to be involved natural language processing. Furthermore, Diffusion Tensor Imaging (DTI) data were acquired to assess differences in structural connectivity between participants that could be related to grammar learning ability. The extent to which participants were able to acquire the rules of the grammar they were exposed to was measured using a grammatical judgement task which took place in the scanner after the second rs-fMRI session. A detailed description of the stimuli and the procedures of both the exposure phase and the grammatical judgement task can be found in Chapter 5.

For each participant, two rs-fMRI sessions (single-shot GE-EPI, 160 repetitions containing 38 transverse slices, TR/TE: 2200/30 ms, voxel size: 2.75 mm$^3$ including a 10% inter-slice gap, FOV: 220 x 217.25 x 104.5 mm) were acquired (before and after exposure). During the 8 min resting state scans, participants were instructed to lay with their eyes closed and to relax, but to not fall asleep. Furthermore, a gradient-echo diffusion weighted scan (60 slices, slice thickness: 2.1 mm, FOV: 224 x 126 x 224 mm) was obtained with diffusion gradients applied in 33 directions with a maximum $b$ value of 1000 s/mm$^2$ and one image with $b = 0$ s/mm$^2$ (60 slices, TR/TE: 6250/70 ms, voxel size: 2.07 x 2.12 x 2.1 mm with no inter-slice gap, FOV: 224 x 224 x 126 mm).

In addition to the rs-fMRI and DTI scans, a high-resolution 3D $T_1$ weighted anatomical scan (140 slices, TR/TE: 9717/459 ms, flip angle: 80°, voxel size: 0.875 x 0.875 x 1.2 mm, FOV: 224 x 168 x 177.3 mm) was acquired for each participant to allow for a good-quality registration between participants and normalization of all participant data to the Montreal Neurological Institute (MNI) template.

### 6.2.4   Resting State fMRI Data analysis

Pre- and post-exposure resting state fMRI data were preprocessed and further analyzed for ROI-to-ROI functional connectivity using the MatLab based CONN functional connectivity toolbox (Whitfield-Gabrieli & Nieto-Castanon, 2012).

The first ten images of each rs-fMRI session were discarded to allow for $T_1$ equilibration. The remaining images were then band-pass filtered over a low frequency window of interest (0,008-0,09 Hz), removing physiological noise and scanner magnetic field drift. They were then detrended, despiked and convolved with a gaussian kernel. Confounding signal from white matter and CSF (five dimensions) as well as movement (six dimensions) and session effects (one dimension per session) were regressed out. Global signal was not removed, because a large portion of the global signal originates from neuronal activity (Schölvinck, Maier, Frank, Duyn, & Leopold, 2010).

Separate conditions were defined for the pre- and post-exposure sessions. Except the one representing Wernicke's area, all seed ROI's selected for first-level analysis were based on the Brodmann atlas included in the CONN toolbox. Based on the large body of research concerning the network involved in language processing, we included left hemisphere (LH) Brodmann's areas (BA) 44 and 45 (Broca's area), LH BA 39 and 40 (angular gyrus and supramarginal gyrus) and LH BA 22 (superior temporal gyrus). To seed Wernicke's area, which spans a number of BA's (Tomasi & Volkow, 2012), we placed a 10 mm spherical ROI at MNI coördinates -51 -51 30 (based on Tomasi & Volkow, 2012). Our analyses were focused on left Pars Triangularis of the Inferior Frontal Gyrus (BA 45), because this area has proven to be important for syntactic processing or (for a review, see Friederici & Gierhan, 2013). Bivariate ROI-to-voxel correlations for the pre- and post-exposure conditions were then computed based on the General Linear Model.

Correlation maps for each participant, session and ROI were then entered into a second-level random-effects analysis to test for functional connectivity on the group level. Second-level covariates were defined for age, sex and behavioral measures obtained during the grammatical judgement task and behavioral post-tests. Because the number of rejected violations during the grammatical judgement task was shown to correlate with individual differences in neural activation during listening (see Chapter 5) regression analyses focused on this measure. Pre- and post-exposure networks for the seeds in Broca's and Wernicke's area were computed and further regression analyses

were masked with the sum of these networks. All functional connectivity results discussed in the results section are computed two-sided and reported as FWE-corrected for multiple comparisons (based on Gaussian Random Field Theory, Worsley et al., 1996) on the cluster level with FWE-corrected clusters $p < 0.05$ considered statistically significant. Only voxels with an uncorrected $p$-value $< 0.001$ were included in the cluster statistics.

### 6.2.5 DTI Data analysis

Diffusion-weighted images were analyzed using the Diffusion II Toolbox implemented in SPM 8 (Statistical Parametric Mapping, version 8, Wellcome Trust Centre for Neuroimaging, London, UK; http://www.fil.ion.ucl.ac.uk/spm/).

The images were first corrected for head movements and eddy currents using rigid-body realignment and full affine co-registration to the $b$0-image. To allow for accurate localization of white matter tracts, the high-resolution $T_1$-weighted anatomical scan of each participant was normalized to the Montreal Neurological Institute (MNI) $T_1$-template. Each participant's diffusion weighted images were then co-registered to the high-resolution anatomical scan and transformed to the MNI template space using the parameters obtained from the normalization of the high-resolution anatomical scan. Images were resampled to 2 mm$^3$ isotropic voxels. After all registration steps were completed, the diffusion gradients were reoriented according to the realignment and normalization parameters.

The diffusion tensor was computed by entering the gradient directions and strengths into a multiple regression model. Diffusion indices Fractional Anisotropy (FA) and Mean Diffusivity (MD) were then computed from the tensor. FA is computed by equation 6.1 (see Le Bihan et al., 2001) and represents the orientation coherence of the diffusion of water protons within a voxel. FA is widely used as a measure of white matter integrity.

$$FA = \frac{\sqrt{3((\lambda_1 - MD)^2 + (\lambda_2 - MD)^2 + (\lambda_3 - MD)^2)}}{\sqrt{2(\lambda_1^2 + \lambda_2^2 + \lambda_3^2)}} \tag{6.1}$$

Linear regression analysis was employed to assess potential correlations between FA and behavioral parameters obtained from the grammatical judgement task. In this analysis, only voxels within the brain's white matter were considered. Explore DTI software (Leemans, Jeurissen, Sijbers, & Jones, 2009) was used to compute FEFA maps where FA is color-coded for the principal diffusion direction, thus allowing for localization on the correlation results in the correct white matter tract. Visualization of diffusion ellipsoids (showing diffusion strength in all directions as well as the principal diffusion direction) on the FEFA maps was also done in ExploreDTI.

## 6.3 Results
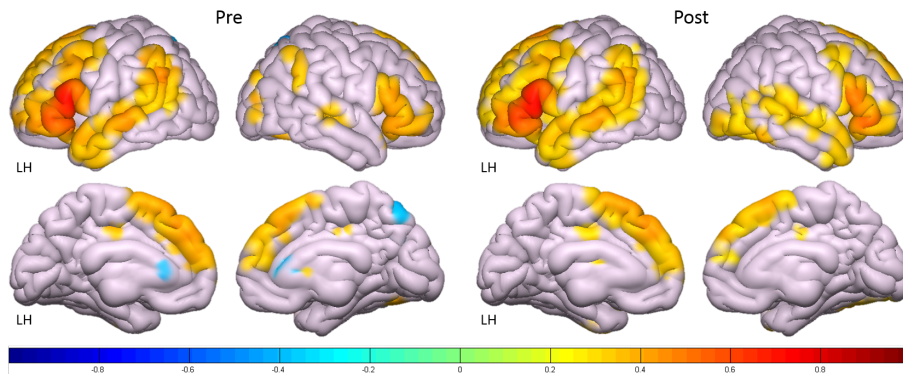


**Figure 6.1: Functional connectivity of Broca's area pre and post exposure to an artificial grammar** Significant ROI-to-voxel correlations at cluster-level $p < 0.05$ FWE-correction. IFG Pars Triangularis is used as an ROI to represent Broca's area. Broca's area is functionally connected to other regions in the left and right Inferior Frontal Gyri, the left and parts of the right Superior Temporal Gyri and Angular Gyri. The pre and post exposure networks did not differ significantly.

### 6.3.1 Behavioral results

The regression analyses conducted in the present study used behavioral parameters from and artificial grammar learning (AGL) task as covariates in order to gain insight into the relation between artificial grammar learning and Resting state functional connectivity and structural connectivity measures.

The AGL task which was used to obtain artificial grammar learning measures is described in detail in Chapter 5. Although participants were generally not able to fully acquire the artificial grammar, results showed large intersubject variability in the success with which they gave grammatical judgements on grammatical and ungrammatical stimuli after auditory exposure to the artificial grammar (% correct responses, mean (sd): 48% (13%)). They were also better able to correctly reject violations of the grammar than to accept grammatical stimuli (two-tailed paired t-test: $T$ = -3.1, $p$ = 0.005).

### 6.3.2 Functional connectivity

Both before and after auditory exposure to the artificial grammar, a reliable functional connectivity was observed between Broca's area (seed ROI: LH Tri) and other area's involved in language processing Figure (6.1). Paired t-tests comparing the pre- and post-exposure networks did not show any significant differences.

| Cond | MNI coördinates | | | Vox. | L/R | Brain region |
|------|-----|-----|-----|------|-----|--------------|
|      | x | y | z |      |     |              |
| Pre  | -56 | 22 | 18 | 2164 | L | Inferior Frontal Gyrus |
|      | -54 | -56 | 28 | 693 | L | Supramarginal / Ang Gyrus |
|      | -42 | 10 | 46 | 573 | L | Premotor cortex / DFC |
|      | -52 | 10 | -20 | 293 | L | MTG / Temporal Pole |
|      | -64 | -30 | -2 | 521 | L | STG / MTG |
|      | 58 | 22 | 0 | 694 | R | Inferior Frontal Gyrus |
|      | 50 | -18 | -8 | 151 | R | STG / MTG |
|      | 0 | 44 | 40 | 1650 | L/R | Medial Frontal Cortex |
|      | -4 | 48 | 16 | 246 | L/R | An PFC / DLPFC |
| Post | -56 | 20 | 16 | 7262 | L | Inferior Frontal Gyrus |
|      |     |     |     |     |     | STG / MTG / Temp. Pole |
|      |     |     |     |     |     | Supramarginal / Ang Gyrus |
|      |     |     |     |     |     | PMC / An PFC / DLPFC / DFC |
|      |     |     |     |     |     | Insula |
|      | -12 | 16 | 64 | 3680 | L/R | Medial Frontal Cortex |
|      | 50 | 32 | -8 | 1918 | R | IFG / PMC / DLPFC |
|      | 48 | -34 | 2 | 720 | R | STG / MTG |
|      | 52 | 10 | -28 | 73 | R | MTG / Temporal Pole |

**Table 6.1: Functional connectivity of Tri in Broca's area prior to and following auditory exposure to an artificial grammar** Values are FWE-corrected with a < 0.0001 threshold, resulting in a minimum cluster extent of 84 voxels and a minimum T-value of 3.88. Cond = Session pre or post exposure to an artificial grammar, L/R = Left/Right hemisphere, Vox. = cluster extent in voxels, Ang. = Angular, DFC = Dorsal Frontal Cortex, MTG = Middle Temporal Gyrus, STG = Superior Temporal Gyrus, (An) PFC = (Anterior) Prefrontal Cortex, DLPFC = Dorsolateral Prefrontal Cortex, PMC = Premotor Cortex, IFG = Inferior Frontal Gyrus

| Brain region | BA | L/R | MNI coördinates | | | $p$-value | Vox. |
|--------------|-----|-----|-----|-----|-----|-----------|------|
|              |     |     | x | y | z |           |      |
| Supramarginal Gyrus | 40 | L | -56 | -52 | 38 | 0.047 | 22 |
| Post. Cingulate Cortex | 30,23 | L | -8 | -64 | 4 | <0.001 | 49 |
| Post. Cingulate Cortex | 31 | R | 8 | -66 | 22 | <0.001 | 58 |
| Sup. Temp / Prim. Aud. | 41,42,22 | L | -60 | -28 | 8 | 0.003 | 37 |
| Sup. Temporal Gyrus | 22 | R | 50 | -24 | 0 | 0.031 | 24 |

**Table 6.2: Regions of which the functional connectivity with Tri correlated with artificial grammar learning** Values are FWE-corrected with a $p < 0.05$ threshold (cluster extents > 22 voxels). BA = Brodmann area, L/R = Left/Right hemisphere, Vox. = cluster extent in voxels, Post. = posterior, Sup. = superior, Temp = Temporal, Prim. Aud. = Primary Auditory Cortex.

In order to assess the role of functional connectivity within the language network in artificial grammar learning, we tested for potential correlations between individual differences in artificial grammar learning and functional connectivity within the network found to be functionally connected to Broca's
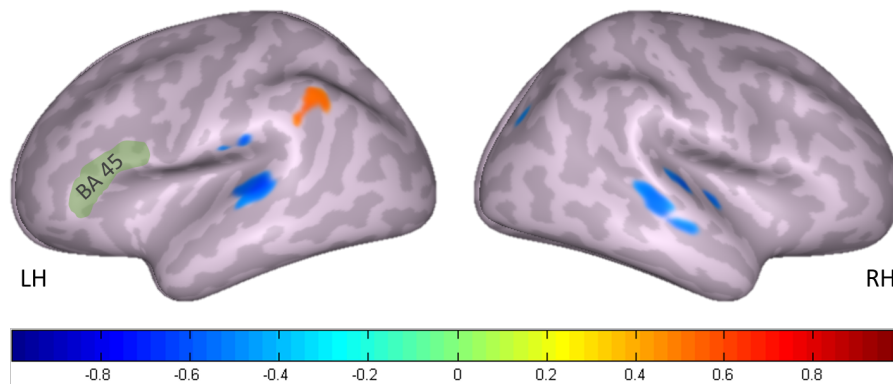
**Figure 6.2: Area's where functional connectivity of Broca's area correlates with artificial grammar learning success after exposure** Correlation map shows significant correlations at cluster-level with $p < 0.05$ FWE-correction. IFG Pars Triangularis (BA 45, shown in green) is used as a seed ROI in Broca's area. The left Supramarginal Gyrus (BA 40) shows a correlation with positive functional connectivity, while the bilateral superior temporal gyri show a correlation with negative functional connectivity. Hemispheres are inflated to show the negative correlation in the right Insula and Superior Temporal Sulcus

area prior to and following auditory exposure to an artificial grammar. Because earlier results showed that individual differences in the ability to reject violations after exposure to an artificial grammar correlate with brain activation during exposure (see Chapter 5), we used linear regression to test for correlations between this parameter of artificial grammar learning and functional connectivity. Our results showed that following, but not prior to auditory exposure to an artificial grammar, functional connectivity between Pars Triangularis and a cluster in the left Supramarginal Gyrus (22 voxels, $p = 0.047$) correlated with the artificial grammar learning performance of participants. Furthermore, negative correlations between artificial grammar learning and functional connectivity were found in the bilateral Posterior Cingulate Cortex (Left: 58 voxels, $p < 0.001$; Right: 49 voxels, $p < 0.001$), and in two clusters covering part of the left Primary Auditory Cortex and Superior Temporal Gyrus (37 voxels, $p = 0.003$; 24 voxels, $p = 0.031$).

A comparison of the correlation maps (Figure 6.2) shows that both the Supramarginal Gyrus cluster and the STG clusters are part of the network that is functionally connected with Broca's area on the group level both before and after auditory exposure to an artificial grammar. Therefore, the negative correlation in the bilateral STG cluster reflects less but present functional connectivity between IFG and STG in more successful learners of the artificial grammar.
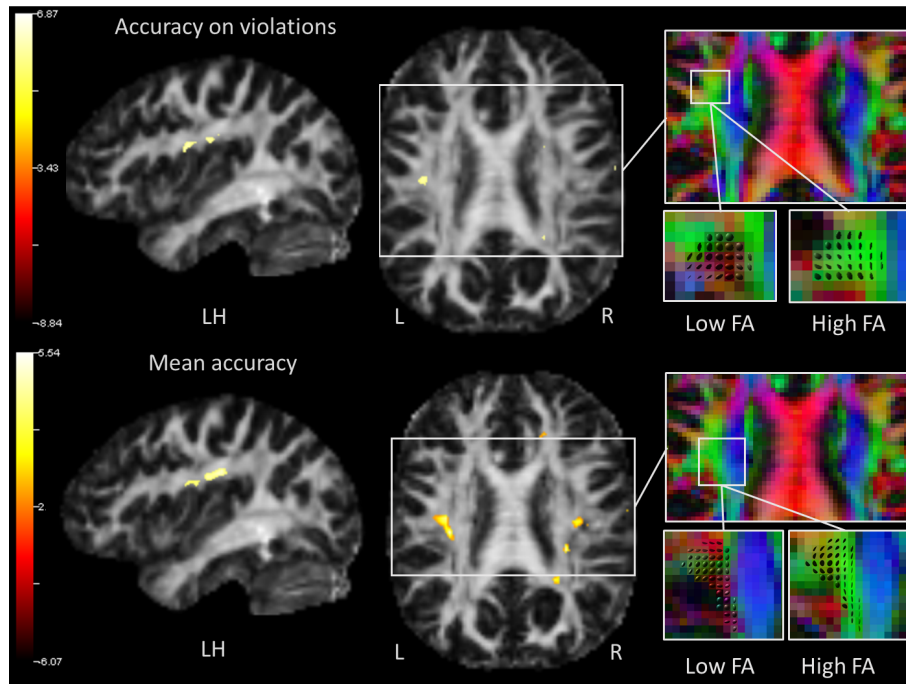
**Figure 6.3: Correlation between artificial grammar learning and white matter integrity in the Arcuate Fasciculus** Saggittal and Axial FA maps with correlation clusters are shown on the left and FEFA maps with diffusion ellipsoids for participants showing high and low FA's in the correlation clusters are shown on the right. Color-code principal diffusion direction: green = anterior-posterior, blue = dorsal-ventral, red = left-right. Top: Correlation between accuracy on violations and FA. Bottom: Correlation between mean accuracy and FA.

### 6.3.3   Structural connectivity

Linear regression analyses using both the accuracy on violations and mean accuracy on the grammatical judgement task as regressors showed correlations between fractional anisotropy in the Arcuate Fasciculus and both measures of artificial grammar learning.

In order to confirm the localization of the correlations within the Arcuate Fasciculus, clusters where a correlation with either mean accuracy or accuracy on violations was found were overlaid on an FA map which was then visually compared with the FEFA map which is color-coded for main diffusion direction. A comparison between these maps and a Diffusion Tensor Imaging atlas (Catani & Thiebaut de Schotten, 2008) showed that the correlation clusters were located in the fronto-parietal and temporal portions of the Arcuate Fasciculus.

The success with which participants rejected violations after exposure to an artificial grammar correlated with FA in the fronto-parietal portion of the

left Arcuate Fasciculus and ($p$ = 0.002, FWE-corrected for whole-brain white matter: Figure 6.3 and Table 6.3). Furthermore, FA in the bilateral fronto-parietal and right temporal Arcuate Fasciculus correlated with the mean accuracy on grammaticality judgements made on grammatical stimuli and violations together (Left: $p < 0.001$; Right: $p < 0.001$, FWE-corrected for whole-brain white matter: Figure 6.3 and Table 6.3). In order to test for group differences, an independent samples t-test was conducted on two groups selected based on a median split on artificial grammar learning scores. This analysis showed no significant differences (Left Occipital-Frontal Fasciculus, MNI cöordinates -34 -44 0: $T$max = 5.37, $p$ = 0.063). Mean diffusivity (MD) was not found to be correlated with any of the behavioral measures.

Although no correlation was found between age and any of the behavioral parameters of artificial grammar learning, previous research has shown a decline in white matter integrity with age (for a review, see Moseley, 2002). A regression analysis testing for the possible negative correlation between age and FA only revealed a correlation in the left Frontal-Occipital Fasciculus (MNI cöordinates: -32 -40 14, $R^2$max = 0.60, $p$ = 0.042, FWE-corrected for white matter). No negative correlation of age with MD was found within the brain's white matter. This suggests that the relation between individual differences in artificial grammar learning and white-matter integrity is not driven by age-related differences in FA.

## 6.4 Discussion

The present study aimed to investigate how individual differences in artificial grammar learning capacity are related to functional and structural connectivity in the healthy adult brain. Individual differences in artificial grammar learning were shown to be reflected in both functional connectivity of the frontal part of Broca's area, to other regions within the language network and white matter integrity in the Arcuate Fasciculus (AF).

| Measure | AF section | L/R | MNI cöordinates | | | FA | $T$max | $R^2$max |
| | | | x | y | z | | | |
|---|---|---|---|---|---|---|---|---|
| Viol. acc | Front-Par | L | -40 | -14 | 24 | 0.65 | 4.67 | 0.55 |
| Mean acc | Front-Par | L | -34 | -22 | -24 | 0.71 | 6.82 | 0.72 |
| | | R | 40 | -22 | -28 | 0.66 | 5.02 | 0.58 |
| | Temp | R | 36 | -30 | 2 | 0.80 | 6.16 | 0.68 |

**Table 6.3: Clusters with a correlation between FA and artificial grammar learning** AF section = section of the Arcuate Fasciculus where the cluster is located, L/R = Left/Right hemisphere, FA = mean Fractional Anisotropy in given cluster, Vox. = cluster extent in voxels, Front-Par = Fronto-Parietal, Temp = Temporal.

### 6.4.1   AGL influences resting-state functional connectivity

Resting-state functional connectivity measurements showed a similar network of brain regions that were functionally connected with IFG Pars Triangularis (Tri), the frontal part of Broca's area before and after auditory exposure to a new artificial grammar. Although AGL did not show a significant effect on the network connected with Tri, behavioral parameters of artificial grammar learning were correlated with functional connectivity after, but not prior to exposure to the grammar. Specifically, after AGL a correlation was found between artificial grammar learning success and functional connectivity between Tri and the Supramarginal Gyrus (SMG). Furthermore, a negative correlation was found between AGL success and functional connectivity between Tri and the bilateral posterior Superior Temporal Gyrus (STG) and Superior Temporal Sulcus (STS). This finding suggests that functional connectivity within the language network is influenced by auditory exposure to an artificial grammar.

The network of brain regions found to be functionally connected to Broca's area after AGL exposure is largely consistent with functional connectivity of Broca's area reviewed in (Tomasi & Volkow, 2012) and fMRI results reported in (C. J. Price, 2010) and includes the bilateral posterior STG and STS. Therefore, the negative correlation between functional connectivity of Broca's area with bilateral STG and STS and AGL success should be interpreted as a lower, not negative, functional connectivity in more successful learners compared to less successful learners after exposure to an artificial grammar. While showing lower functional connectivity between Broca's area and bilateral STG/STS, these participants showed higher functional connectivity between Tri and the Supramarginal Gyrus. This suggests that a relatively strong functional connectivity between Tri area and SMG combined with a weaker FC between Tri and posterior STG might enable participants to learn a grammar more successfully or, alternatively, might be induced by successful artificial grammar learning.

Bilateral posterior STG/STS have been implicated in phonological processes of speech recognition (e.g. Hickok, 2009) and semantic processing (Friederici, Rueschemeyer, Hahne, & Fiebach, 2003). Furthermore, the SMG cluster is adjacent to Wernicke's area and is consistent with the SMG cluster that was found to be functionally connected with Broca's area in a meta-analysis of RS-fMRI data from a large sample of healthy participants (Tomasi & Volkow, 2012). Furthermore, SMG was previously shown to be functionally connected to IFG during speech processing tasks with high processing loads when stimuli were predictable (Obleser, Wise, Dresner, & Scott, 2007). Because the current experiment was conducted under difficult listening conditions and stimuli were more predictable for more successful learners, this might explain the correlation between functional connectivity and artificial grammar learning. Our data suggest that a higher functional connectivity between IFG Pars Triangularis and SMG through the anterior part of the dorsal language path-

way (Hickok & Poeppel, 2004; Parker et al., 2005; Rauschecker & Scott, 2009) aids syntactic processing.

## 6.4.2   Contribution of the dorsal language pathway to AGL

Individual differences in artificial grammar learning capacities were not only related to functional, but also to structural connectivity measures. Our results showed correlations between Fractional Anisotropy in the fronto-parietal section of left Arcuate Fasciculus and the success with which participants rejected violations of a learned artificial grammar, while response accuracy on both grammatical and ungrammatical stimuli correlated with FA in the bilateral fronto-parietal Arcuate Fasciculus.

The Arcuate Fasciculus connects the IFG, Middle Frontal Gyrus and Precentral Gyrus with the Superior and Middle Temporal Lobes and the Supramarginal Gyrus (e.g Catani & Thiebaut de Schotten, 2008; Parker et al., 2005) and has since long been established as an important connection in the language network (Geschwind, 1965). Together with the Superior Longitudinal Fasciculus, the Arcuate Fasciculus forms the anatomical basis for the dorsal language pathway (Hickok & Poeppel, 2004; Saur et al., 2008). Decreased FA in the Arcuate Fasciculus has been shown to play a role in speech and language deficits like stuttering (Chang, Erickson, Ambrose, Hasegawa-Johnson, & Ludlow, 2008) and Specific Language Impairment (Vallée et al., 2014).

The finding that participants who performed better when learning an artificial grammar showed a higher FA in the Arcuate Fasciculus, suggests that the AF and the dorsal language pathway contribute to syntactic processing and language learning. This is in accordance with previous studies showing that white matter integrity in the left (Flöel et al., 2009; Lopez-Barroso et al., 2011) and right (Loui, Li, & Schlaug, 2011) hemisphere can predict different aspects of artificial grammar learning and second language learning (Li, Legault, & Litcofsky, 2014) success.

Purely based on the results of the current study, it cannot be determined whether the presence of more coherent white matter in AF leads to better grammar learning capacities, whether increased grammar learning experience and ability leads to higher FA in the Arcuate Fasciculus through a "use-it-or-lose-it"-effect or whether a third factor drives both. FA can reflect both macro-structural (fiber orientation coherence) and micro-structural (myelination) attributes of white matter and myelination has been suggested as a source of adult plasticity because it extends into adulthood and has been associated with activity-driven plasticity (Fields, 2005). Therefore, increased FA in the Arcuate Fasciculus can be both a cause and an effect of higher (artificial) grammar learning capacity. Further studies addressing the development of individual differences in grammar learning ability will be needed to shed light on this issue.

### 6.4.3   Conclusion
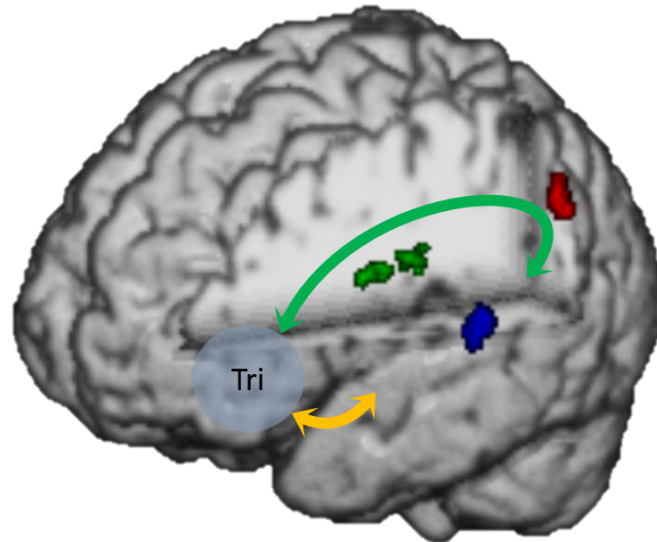


**Figure 6.4: Ventral and dorsal language pathways in relation to the correlations found between functional connectivity and AGL.** Correlation between accuracy on violations and FA in the left Arcuate Fasciculus (green cluster) and positive (red cluster) and negative (blue cluster) correlations between accuracy on violations and FC with Tri. Pathways: green = Dorsal Pathway, yellow = Ventral Pathway. The dorsal pathway runs through the Arcuate Fasciculus and the Superior Longitudinal Fasciculus and the ventral pathway runs through the Extreme Capsule

An integrated analysis of functional and structural connectivity within participants might allow for a better understanding of the underlying mechanisms leading to individual differences in artificial and possibly natural grammar learning. Figure 6.4 summarizes the correlations found between behavior and both functional and structural connectivity. Prior research has shown that artificial grammar learning in older adults is linked to functional and structural connectivity (Antonenko, Meinzer, Lindenberg, Witte, & Flöel, 2012). The present study revealed both a higher FA in the anterior part of the Arcuate Fasciculus, and a higher functional connectivity between Pars Triangularis and the Supramarginal Gyrus in participants who performed better on an artificial grammar learning task. These results suggest that increased connectivity between Tri and SMG through the dorsal language pathway, specifically the Arcuate Fasciculus, is beneficial to artificial grammar learning and might thus play a role in syntactic processing.

It is conceivable that a higher white matter integrity or myelination of AF in more successful learners also allows for increased functional connectivity in these participants. Although the correlation between functional connectivity and artificial grammar learning measures was only found after auditory

exposure to and artificial grammar, it is unlikely that this correlation is purely task-related, because the structural connectivity data show a similar effect. A possible explanation is that successful learners possess a language network with stronger white matter connections, within which important connections are engaged when they are faced with the task of acquiring an artificial grammar. In learners who have more difficulties learning the grammar, the structural connectivity between Tri and SMG might be weaker, allowing for less functional connectivity between these regions.

CHAPTER 7

## Summary and conclusions

*Vocal learning is a rare trait in the animal kingdom, but it is not unique to humans. It is also a trait that is restricted by the brain rather than by a species' vocal apparatus. The present thesis delved into the neural substrates of vocal learning in the two species in which vocal learning behavior has been most extensively studied: zebra finches and humans. Functional Magnetic Resonance Imaging was employed in both species in order to explore potential common neural mechanisms which underlie both the ability to develop species-specific vocalizations through memorization and imitation of an adult model and the restrictions that sensitive periods for vocal learning, which have been hypothesized for both songbirds and humans, impose on this ability. In zebra finches, a songbird species that learns one specific song, neural responses to songs learned during development were compared to neural responses to songs that were familiar, but not learned for production. In human adults, who possess a language system too extensive and complex to test in a single experiment, neural responses were recorded before, during and after learning an artificial grammar, comparing a grammar with and without learnable regularities in the form of non-adjacent dependencies. In Section 7.1, each chapter is briefly summarized and the main findings from each experiment are reviewed. Furthermore, a number of limitations to the comparative approach employed in this thesis are discussed in Section 7.2 and in Section 7.4 suggestions are made for methodological advancements and future studies that might shed light on some of the issues that could not yet be resolved in the research covered by this thesis. To conclude, the implication of our studies for potential common neural substrates of vocal learning as well as the implications for the use of songbird and human fMRI in comparative studies into the neural substrates of vocal learning are stated in Section 7.3.*

# 7.1 Summary

### 7.1.1 Processing of learned song in the auditory midbrain of adult and juvenile zebra finches

In the first part of this thesis functional MRI is applied in the zebra finch model in order to gain insight into the neural basis of vocal learning in songbirds. Chapter 1 describes the development of songbird fMRI, its application for the study of vocal learning and refinements of the method employed in recent studies. Among others these include the detection of subtle BOLD differences induced by songs that are acoustically close and the detection of an auditory BOLD response in juvenile zebra finches, developments which allowed for the studies reported in Chapter 3 and Chapter 4 of this thesis.

Early in the song learning process, zebra finches form a memory trace based on the song of one or several tutors (usually their father), which guides their vocal practice in later stages of song learning. In Chapter 3 and Chapter 4 on-line processing of song in the zebra finch brain is studied using fMRI in order to explore where in the songbird brain this memory trace is formed and how the zebra finch brain processes the tutor song compared to other song stimuli.

Chapter 3 shows that in adult zebra finches the right Dorsal part of the Lateral Mesencephalic nucleus (MLd), the main auditory midbrain nucleus in the songbird brain, selectively responds to the song of the bird's tutor. This response is shown to be distinct from the selective response to the bird's own song, which is also found in right MLd (Poirier et al. (2009) and Chapter 3 of this thesis), and does not reflect acoustic differences or familiarity differences between the stimuli. Furthermore, a positive correlation was observed between the amplitude of the differential activation induced by the tutor song and the strength of song learning, as expressed by the similarity of a bird's song to the song of its tutor. These findings suggest that early auditory experience creates a lasting memory trace of the adult song model in the auditory midbrain of the zebra finch, which might subserve further stages of song learning.

The aim of the study described in Chapter 4 was twofold. Firstly, this study aimed to establish functional MRI in juvenile zebra finches, which was attempted for the first time in this study. Secondly, this study aimed to explore the development of the selectivity for the adult song model revealed in Chapter 3, in relation to song learning. Taking a longitudinal approach, fMRI data were collected from juvenile zebra finches at different stages during their song development. Results showed a strong selectivity for the adult song model in the left MLd of male zebra finches at 60 days post hatching (DPH). In contrast, at 500 DPH, the same individuals show tutor song selectivity in right MLd, confirming the findings of Chapter 3. The auditory midbrain of male zebra finches was thus demonstrated to develop selectivity for the adult song

model towards the end of song memorization, but the lateralization of this selectivity changes after song learning is completed. Female zebra finches do not learn to sing, but the males' song plays an important role in their mate choice. It was thus not surprising that the left MLd of 30 DPH juvenile female zebra finches only showed selectivity for conspecific over heterospecific song. At later ages, the female selectivity for conspecific song was no longer significant. This early sensitivity to conspecific song in female birds might contribute to the selection of quality mates in adulthood (Lauay et al., 2004).

### 7.1.2 Artificial grammar learning in human adults

In the second part of this thesis, human fMRI studies are discussed which were designed to shed light on the neural substrates of our own capacity for vocal learning, the human language learning capacity. However, learning a human language encompasses an intricate process of which the resulting linguistic knowledge of even an individual speaker cannot be captured in a single stimulus or set of stimuli. The human language learning capacity is characterized by the ability to derive the presence and the nature of syntactic relations between words and constituents from the input and use this knowledge to express meaning in an uncontrained way, despite to constrained number of words and sentences that a single human can memorize. Because this capacity to learn syntactic rules is so central to language and is also an ability hypothesized to be most impacted by the closing of the sensitive period for language acquisition, we chose to simulate it in adult humans and compare its neural correlates to those of tutor song learning in zebra finches. In order to investigate the neural correlates of the acquisition of syntactic rules, Chapters 5 and 6 employed fMRI, resting state functional connectivity and Diffusion Tensor Imaging (DTI) in an Artificial Grammar Learning (AGL) paradigm with a grammar containing non-adjacent dependencies.

In Chapter 5 the brain responses of adult subjects are examined who are faced with the task (i) to learn an artificial grammar containing non-adjacent dependencies from mere auditory exposure and (ii) to subsequently give grammaticality judgments on grammatical and ungrammatical stimuli. Although in the limited time of exposure within the experiment most participants were not able to learn the language to such a degree that they performed above chance level on the grammatical judgment task, fMRI results showed that they did develop a neural sensitivity to the non-adjacent dependencies that were present in the artificial grammar.

FMRI data from the auditory exposure of participants to the grammar containing non-adjacent dependencies (NAD grammar) and a control grammar showed a correlation between the extent to which individual participants showed sensitivity to the non-adjacent dependencies and the amplitude of the differential activation in response to the NAD grammar in the left Inferior Frontal Gyrus (LIFG) and bilateral Superior Temporal Gyri and Insulae. Furthermore, when giving grammaticality judgments, participants showed a

differential activation in LIFG in response to correctly judged compared to incorrectly judged items only when judging ungrammatical stimuli. Additionally, participants were shown to be more sensitive to ungrammatical stimuli, both at the behavioral and at the neural level. These results suggest that neural activity in language-related brain regions during learning reflects the degree to which grammatical rules are derived from the input. This activation might represent the formation of a syntactic rule in memory, leading a high sensitivity to violations of this rule.

The relation between learning and neural activation of language-related brain regions was further supported by the results from connectivity measures discussed in Chapter 6, which were obtained during the same artificial grammar learning study. These data showed that functional as well as structural connectivity between part of Broca's area and primarily left superior temporal and inferior parietal regions in the language network was related to individual differences in artificial grammar learning and possibly also to their natural language learning capacity. Functional connectivity data from resting state fMRI scans showed a correlation between frontal-parietal connectivity in the left hemisphere, possibly through the dorsal language pathway, and artificial grammar learning. Moreover, Fractional Anisotropy, a measure of white matter integrity, in the left Arcuate Fasciculus was shown to correlate with the degree to which participants had learned the grammar. These results suggest that structural and functional connectivity along the dorsal language pathway might subserve the capacity for implicit learning of artificial grammars and possibly natural language syntax.

## 7.2 Discussion

### 7.2.1 Implications for neural correlates of vocal learning

Adult as well as juvenile zebra finch fMRI data show that MLd, the zebra finch auditory midbrain, differentiates the tutor's song from other conspecific songs. Furthermore, the strength of the selectivity is related to the strength of song learning in adult birds. These results suggest that MLd plays a role in tutor song memorization.

Furthermore, lateralization of tutor song selectivity was found to be relatively unstable in juvenile birds with a clear left lateralization at 60 DPH which was no longer observed at 100 DPH. However, in adulthood, male zebra finch from both studies showed a stable right lateralization of selective activation for the tutor song. These findings suggest that the final lateralization of tutor song selectivity is reached after song learning, implying that lateralization of selective responses in the midbrain, like in telencephalic regions (Phan & Vicario, 2010), depends on auditory experience. Exposure to a tutor during the sensory phase thus creates a memory trace of the tutor's song which is more specific in more successful learners and becomes right lateralized after

prolonged auditory experience.

The timing of the stabilization of lateralization with auditory experience coincides with song crystallization and the end of the sensitive period for vocal learning. This neural mechanism might underlie the changes in the nature and function of tutor song memory during development. During song learning, tutor song memory is plastic and continuously shapes song production, while it is no longer plastic and only needed for song maintenance in adulthood.

Our artificial grammar learning studies in adult human participants also revealed a correlation between behavioral learning measures and selective activation for the learned grammar in brain regions implicated in language processing, among which the inferior frontal gyrus, which has been associated with processing of syntactic rules (Petersson et al., 2012) and was hypothesized to show a gradient of abstraction (Uddén & Bahlmann, 2012). When judging new phrases on grammaticality, this region was also selectively activated in response to ungrammatical stimuli that were correctly rejected compared to ungrammatical stimuli that were accepted. These results confirm the role of LIFG or Broca's area in artificial grammar learning, which has been proven to be a good model for natural syntax learning (Petersson et al., 2012) and suggest that individual differences in neural activation during learning are related to individual differences in language learning outcome. The absence of a correlation with age suggests that the individual differences in learning capacity and selective activation are not caused by a sensitive period effect, but might represent differences in language learning capacity rooted in development, but independent of age in adulthood. Thus, in human participants both on-line artificial grammar learning (Chapter 5) and processing of learned linguistic structures in artificial grammar as well as natural language (Petersson et al., 2012) selectively activates brain regions associated with auditory processing and more importantly, left lateralized brain regions which process stimuli with a high level of abstraction (Uddén and Bahlmann (2012) and Chapter 5 of this thesis).

Furthermore, additional assessment of structural and functional connectivity uncovered a relation between artificial grammar learning capacity and functional as well as structural connectivity between Broca's area and temporal and parietal regions in the left hemisphere. Although white matter changes greatly surpass the time-scale of a single fMRI session, anisotropy does increase with age in childhood and adolescence (Barnea-Goraly et al., 2005; Schmithorst, Wilke, Dardzinski, & Holland, 2002) and can be enhanced by training in adulthood (e.g Scholz, Klein, Behrens, & Johansen-Berg, 2009). This suggests that the individual differences in white matter integrity in the arcuate fasciculus could either be the result of developmental structural differences, in turn leading to differences in grammar learning capacities or that the individual differences in both white matter structure and grammar learning capacities are the result of experience.

In contrast to white matter structure, individual differences in functional connectivity were only related to artificial grammar learning measures after auditory exposure to the artificial grammar. This indicates that functional connectivity within the language network is modulated by artificial grammar learning. The fact that functional connectivity between Broca's area and the inferior parietal lobule (IPL) positively correlates with artificial grammar learning, while functional connectivity between Broca's area and the auditory cortex negatively correlates with artificial grammar learning, supports an interpretation in terms of levels of abstraction in processing. IPL activation during syntactic processing was shown to be related to second language proficiency (Wartenburger et al., 2003), suggesting that participants who learn more successfully engage regions implicated in syntactic processing while less successful learners engage regions that process stimuli at a lower level of abstraction.

### 7.2.2   Insights from cross-species fMRI studies

The studies described in this thesis encompass different species, different ages and show results in different brain regions. However, all of the tasks performed by subjects during fMRI measurements assessed how vocal learning impacts brain function. Together, the studies on both zebra finches and humans showed that, in both species, the brain responds selectively to vocalizations that are learned over those that are acoustically very close, but were not learned. Furthermore, even though in the zebra finch experiment the exposure to the learned stimuli took place during development while the human adults learned the artificial grammar from mere exposure in the same experiment, the strength of the selective responses was related to behavioral learning strength in both species.

Although these effects are found in different species and in different brain regions, they may indicate how brains process learned vocalizations during adulthood and vocalizations of adult conspecifics during language acquisition or song learning. The presence of selective processing might indicate that a memory is formed, which later can be used for either the perception of vocalization uttered by others or the processing of auditory feedback of one's own vocalizations in order to adapt them. To gain insight in any similarities in the relation between perception and production, the role of this neural selectivity in processing auditory feedback during production should be studied in both species, but the results in this thesis indicate that similar memory traces of learned speech and birdsong might develop through mere exposure.

This thesis aimed to gain insight in both similarities between the neural correlates of birdsong learning and language acquisition and in the value of the songbird model for vocal learning to brain studies of human language learning. Our studies have shown that, while similar stimuli presented to humans and songbirds might induce very different brain responses, the study of a similar behavioral process, such as learning new vocal structures, can uncover similar brain-behavior coupling in two very distantly related species.

For the songbird model of vocal learning this means that it can inform the neural basis of human language acquisition and evolution as long as the behavior is not left out of the analysis. Because of the evolutionary distance between humans and songbirds one should keep in mind that convergence on the behavioral level might go hand in hand with convergence on the neural level, meaning that non-homologous brain regions might subserve similar neural processes underlying vocal learning.

### 7.2.3 Limitations of cross-species fMRI studies

The human and zebra finch fMRI data presented in this thesis showed some similar relations between neural activation and vocal learning, but these effects were not observed in homologous brain regions. Although learning of new speech sounds was previously shown to correlate with activation of the inferior colliculus (IC), the human auditory midbrain region (Chandrasekaran et al., 2012), the role of the human IC in vocal learning is not assessed in our studies. Partly, this is due to technical limitations, including the fact that fMRI in IC is complicated by its adjacency to the fourth ventricle, which is a potential cause of artifacts due to the continuous movement of cerebrospinal fluid through this ventricle.

However, based on earlier AGL studies, the level of abstraction and generalization of auditory information needed for artificial grammar learning is likely to depend on cortical rather than sub-cortical regions (e.g. Folia & Petersson, 2014; Uddén & Bahlmann, 2012; Yang & Li, 2012). In vocal learning, gradients of abstraction as proposed by Uddén and Bahlmann (2012) might not be restricted to the left Inferior Frontal Gyrus, but might also be present in the auditory pathway (Ranasinghe, Vrana, Matney, & Kilgard, 2013). Lacking the productivity of human language, memorization of the tutor's song does not require the same level of abstraction as artificial grammar learning. Therefore, similar neural mechanisms are observed at different levels of abstraction in the human and zebra finch brain.

When discussing levels of abstraction of auditory stimuli it should be kept in mind that the zebra finch brain activation found in Chapters 3 and 4 was measured under anesthesia, while the human artificial grammar learning studies (Chapters 5 and 6) were conducted in awake and attentive participants. Imaging awake animals, however, requires extensive training (e.g. De Groof et al., 2013) and this technique has not yet been developed for zebra finches (see Chapter 2). Although song selectivity in the auditory midbrain of zebra finches is not expected to be markedly influenced by anesthesia, the degree of abstraction which is employed when processing the stimuli might have been influenced. This would explain the lack of selective activation in the Caudomedial Nidopallium (NCM), a secondary auditory area which had previously been linked to tutor song memory (Bolhuis et al., 2000) and might thus complicate the comparison with our human studies, in which participants were awake and the task required a higher level of abstraction.

Another limitation of the studies presented in this thesis is inherent to the use of auditory stimuli, which was done in order to preserve the role of auditory processing in the perception of human speech and birdsong. MRI scanners are noisy environments due to the switching of gradients during scanning. This complicates the delivery of auditory stimuli to participants. For the zebra finch studies, stimuli were delivered by small commercial loudspeakers with the magnets removed (see Chapter 2), resulting in amplification but minimal distortion of the stimuli. For safety reasons, in the human studies stimuli were delivered via pneumatic headphones, which led to minimal distortion and produced a maximum sound level at which participants needed to stay attentive in order to correctly identify the phonemes. Neuroimaging results from both species and reports from the human participants on the discernibility of the auditory stimuli indicated that stimuli were properly perceived and processed. However, the precise influence of the sound quality and perceived loudness on the neural activation in either species cannot be determined.

## 7.3 Conclusions

Despite their limitations, the studies described in this thesis provide additional insight into the neural correlates of vocal learning and the processing of learned stimuli. Both when learning their own vocalizations and processing those of others, humans and songbirds need to abstract the acoustic input at some level. Human language is characterized by a significantly higher level of abstraction than zebra finch song, resulting in distinct patterns of neural activation. In spite of these considerable differences, tutor song processing in zebra finches and artificial grammar learning in humans showed a number of neuro-functional similarities, allowing for the following conclusions:

- In both humans and zebra finches, vocal learning and processing of learned species-specific vocalizations and their structure induce selective activation in brain regions associated with auditory processing and abstraction of vocal signals.

- The amplitude of this selective activation is related to learning outcomes both in humans and in zebra finches

- Structural and functional connectivity within the network processing human language is related to (artificial) grammar learning capacity.

Within the scope of this thesis, not all measures could be obtained in both species or both developmental stages, leaving some questions unanswered regarding the shared neurobiological mechanisms underlying vocal learning. Section 7.4 will touch upon a number of the issues that could not be addressed in this thesis and proposes how future studies could shed light on them.

## 7.4 Future Prospects

The present thesis encompasses a series of studies aimed at uncovering the common neural correlates of vocal learning in songbirds and humans. Songbird studies were able to show how learned vocalizations are processed in the juvenile and adult brain, while the human fMRI studies assessed both the neural correlates of on-line learning and those of processing artificial grammar. Additionally, resting state and DTI data provided insight into the role of neural connectivity in artificial grammar learning. In both species, selective activation for learned vocalizations was found. However, a number of questions still remain.

### 7.4.1 Artificial grammar learning within the sensitive period for language development

In songbirds, neural substrates for tutor song perception were observed to change after song learning. Because our human studies only included adult participants, our data do not allow us to place the neural correlates for grammar learning in a developmental perspective. Currently, additional studies are run, which replicate the same experiment in pre-pubertal children in order to uncover whether and how the neural correlates of artificial and possibly natural grammar learning are influenced by the sensitive period for language development. Both our study assessing tutor song selectivity throughout song development in juvenile zebra finches and infant neuroimaging studies addressing early speech perception (Dehaene-Lambertz, Dehaene, & Hertz-Pannier, 2002) observed a different lateralization of activation induced by learned vocalizations. Therefore, lateralization of activation induced by artificial grammar learning, in particular in the inferior frontal gyrus, is expected to be less profoundly left lateralized in school-age children.

In addition to fMRI measurements during artificial grammar learning, the study run in children, like the one described in adults, includes measurements of functional and structural connectivity, potentially answering the question whether the correlation between white matter integrity and artificial grammar learning performance can be ascribed to developmental or experience-dependent white matter changes. If age-related white matter changes in children are also found to be correlated with artificial grammar learning capacity, the individual differences observed in adult participants are likely to have a developmental origin.

### 7.4.2 Different levels of abstraction

The tasks administered to songbirds and human participants of the studies described in this thesis could be expected not to elicit activation in homologous regions, because the level of abstraction that was needed to learn the

grammar or recognize the learned vocalizations differed greatly. This correlation between level of abstraction and species limits the possibility to draw any conclusions about the neural correlates of the ability of zebra finches to generalize over vocalizations (for a review, see Ten Cate & Okanoya, 2012). Zebra finches were observed to generalize prosodic (Spierings & Ten Cate, 2014) rather than syntactic patterns (Van Heijningen et al., 2009) to new stimuli. Consequently, a task that requires abstraction of frequency patterns might for example recruit the Caudomedial Nidopallium (NCM), a region in the zebra finch auditory pathway that contains more complex neurons than MLd (for a recent review, see Woolley, 2012) and might therefore allow for more complex computations.

Chapter 3 showed that the zebra finch auditory midbrain discriminates between learned and non-learned, though familiar, vocalizations, suggesting that this nucleus exhibits a rudimentary auditory memory function. Because the human inferior colliculus (IC), the homologue of MLd, plays a role in learning tones of foreign (tone) languages (Chandrasekaran et al., 2012), a similar role for IC in the recognition of native phonemes is conceivable. Further fMRI studies focusing exclusively on IC could shed light on the role of this nucleus in the process of discriminating native-language from non-native phonemes in language processing.

Lastly, fMRI studies in awake and attentive zebra finches might enable birds to process vocalizations at a higher level of abstraction, which might allow for studies addressing the neural correlates of rule learning in zebra finches. This technical advancement, together with developments that would introduce a social context, might also enable zebra finches to learn in the scanner, allowing for real-time imaging studies of vocal learning.

# References

Akutagawa, E., & Konishi, M. (2010). New brain pathways found in the vocal control system of a songbird. *Journal of Comparative Neurology*, *518*(15), 3086–3100.

Amaro, E., Williams, S. C., Shergill, S. S., Fu, C. H., MacSweeney, M., Picchioni, M. M., . . . McGuire, P. K. (2002). Acoustic noise and functional magnetic resonance imaging: current strategies and future prospects. *Journal of Magnetic Resonance Imaging*, *16*(5), 497–510.

Amin, N., Doupe, A., & Theunissen, F. E. (2007). Development of selectivity for natural sounds in the songbird auditory forebrain. *Journal of Neurophysiology*, *97*(5), 3517–3531.

Amin, N., Grace, J., & Theunissen, F. (2004). Neural response to birds own song and tutor song in the zebra finch field l and caudal mesopallium. *Journal of Comparative Physiology A*, *190*(6), 469–489.

Andalman, A. S., & Fee, M. S. (2009). A basal ganglia-forebrain circuit in the songbird biases motor output to avoid vocal errors. *Proceedings of the National Academy of Sciences*, *106*(30), 12518–12523.

Antonenko, D., Meinzer, M., Lindenberg, R., Witte, A. V., & Flöel, A. (2012). Grammar learning in older adults is linked to white matter microstructure and functional connectivity. *Neuroimage*, *62*(3), 1667–1674.

Aronov, D., Andalman, A. S., & Fee, M. S. (2008). A specialized forebrain circuit for vocal babbling in the juvenile songbird. *Science*, *320*(5876), 630–634.

Aslin, R. N., & Fiser, J. (2005). Methodological challenges for understanding cognitive development in infants. *Trends in Cognitive Sciences*, *9*(3), 92–98.

Astrakas, L. G., & Argyropoulou, M. I. (2010). Shifting from region of interest (roi) to voxel-based analysis in human brain mapping. *Pediatric Radiology*, *40*(12), 1857–1867.

Baddeley, A. (2003). Working memory and language: an overview. *Journal of Communication Disorders*, *36*(3), 189–208.

Bahlmann, J., Schubotz, R. I., & Friederici, A. D. (2008). Hierarchical artificial grammar processing engages broca's area. *Neuroimage*, *42*(2), 525–534.

Baptista, L. F., & Schuchmann, K.-L. (1990). Song learning in the anna hummingbird (calypte anna). *Ethology*, *84*(1), 15–26.

Barnea-Goraly, N., Menon, V., Eckert, M., Tamm, L., Bammer, R., Karchemskiy, A., . . . Reiss, A. L. (2005). White matter development during childhood and adolescence: a cross-sectional diffusion tensor imaging study. *Cerebral Cortex*, *15*(12), 1848–1854.

Bauer, E. E., Coleman, M. J., Roberts, T. F., Roy, A., Prather, J. F., & Mooney, R. (2008). A synaptic basis for auditory–vocal integration in the songbird. *The Journal of Neuroscience*, *28*(6), 1509–1522.

Beecher, M. D., & Burt, J. M. (2004). The role of social interaction in bird song learning. *Current Directions in Psychological Science*, *13*(6), 224–228.

Birdsong, D., & Molis, M. (2001). On the evidence for maturational constraints in second-language acquisition. *Journal of Memory and Language*, *44*(2), 235–249.

Black, K. J., Snyder, A. Z., Koller, J. M., Gado, M. H., & Perlmutter, J. S. (2001). Template images for nonhuman primate neuroimaging: 1. baboon. *Neuroimage*, *14*(3), 736–743.

Boersma, P. (2002). Praat, a system for doing phonetics by computer. *Glot International*, *5*(9/10), 341–345.

Bolhuis, J. J., & Gahr, M. (2006). Neural mechanisms of birdsong memory. *Nature Reviews Neuroscience*, *7*(5), 347–357.

Bolhuis, J. J., Hetebrij, E., Den Boer-Visser, A. M., De Groot, J. H., & Zijlstra, G. G. (2001). Localized immediate early gene expression related to the strength of song learning in socially reared zebra finches. *European Journal of Neuroscience*, *13*(11), 2165–2170.

Bolhuis, J. J., Okanoya, K., & Scharff, C. (2010). Twitter evolution: converging mechanisms in birdsong and human speech. *Nature Reviews Neuroscience*, *11*(11), 747–759.

Bolhuis, J. J., Zijlstra, G. G., Den Boer-Visser, A. M., & Van der Zee, E. A. (2000). Localized neuronal activation in the zebra finch brain is related to the strength of song learning. *Proceedings of the National Academy of Sciences*, *97*(5), 2282–2285.

Bonatti, L. L., Pena, M., Nespor, M., & Mehler, J. (2005). Linguistic constraints on statistical computations the role of consonants and vowels in continuous speech processing. *Psychological Science*, *16*(6), 451–459.

Booth, J. R., Wood, L., Lu, D., Houk, J. C., & Bitan, T. (2007). The role of the basal ganglia and cerebellum in language processing. *Brain Research*, *1133*, 136–144.

Boumans, T., Gobes, S. M., Poirier, C., Theunissen, F. E., Vandersmissen, L., Pintjens, W., . . . Van der Linden, A. (2008). Functional mri of auditory responses in the zebra finch forebrain reveals a hierarchical organisation based on signal strength but not selectivity. *PloS One*, *3*(9), e3184.

Boumans, T., Theunissen, F. E., Poirier, C., & Van der Linden, A. (2007). Neural representation of spectral and temporal features of song in the auditory forebrain of zebra finches as revealed by functional mri. *European Journal of Neuroscience*, *26*(9), 2613–2626.

Boumans, T., Vignal, C., Smolders, A., Sijbers, J., Verhoye, M., Van Audekerke, J., ... Van der Linden, A. (2008). Functional magnetic resonance imaging in zebra finch discerns the neural substrate involved in segregation of conspecific song from background noise. *Journal of Neurophysiology*, *99*(2), 931.

Braaten, R. F., & Reynolds, K. (1999). Auditory preference for conspecific song in isolation-reared zebra finches. *Animal Behaviour*, *58*(1), 105–111.

Brainard, M. S. (2004). Contributions of the anterior forebrain pathway to vocal plasticity. *Annals of the New York Academy of Sciences*, *1016*(1), 377–394.

Brainard, M. S., & Doupe, A. J. (2000). Interruption of a basal ganglia–forebrain circuit prevents plasticity of learned vocalizations. *Nature*, *404*(6779), 762–766.

Brainard, M. S., & Doupe, A. J. (2002). What songbirds teach us about learning. *Nature*, *417*(6886), 351–358.

Cannestra, A. F., Santori, E. M., Holmes, C. J., & Toga, A. W. (1997). A three-dimensional multimodality brain map of the nemestrina monkey. *Brain Research Bulletin*, *43*(2), 141–148.

Capsius, B., & Leppelsack, H.-J. (1996). Influence of urethane anesthesia on neural processing in the auditory cortex analogue of a songbird. *Hearing Research*, *96*(1), 59–70.

Cardin, J. A., & Schmidt, M. F. (2003). Song system auditory responses are stable and highly tuned during sedation, rapidly modulated and unselective during wakefulness, and suppressed by arousal. *Journal of Neurophysiology*, *90*(5), 2884–2899.

Cardin, J. A., & Schmidt, M. F. (2004). Auditory responses in multiple sensorimotor song system nuclei are co-modulated by behavioral state. *Journal of Neurophysiology*, *91*(5), 2148–2163.

Carpenter, M., & Tomasello, M. (1995). Joint attention and imitative learning in children, chimpanzees, and enculturated chimpanzees*. *Social Development*, *4*(3), 217–237.

Castelino, C. B., & Ball, G. F. (2005). A role for norepinephrine in the regulation of context-dependent zenk expression in male zebra finches (taeniopygia guttata). *European Journal of Neuroscience*, *21*(7), 1962–1972.

Castelino, C. B., & Schmidt, M. F. (2010). What birdsong can teach us about the central noradrenergic system. *Journal of Chemical Neuroanatomy*, *39*(2), 96–111.

Catani, M., & Thiebaut de Schotten, M. (2008). A diffusion tensor imaging tractography atlas for virtual in vivo dissections. *Cortex*, *44*(8), 1105–1132.

Central Committee on Research Involving Human Subjects, C. C. M. O. (2004). *Memorandum: Criteria for assessing non-therapeutic mri scans on minor and incapacitated research subjects.* available at http://www.ccmo.nl/attachments/files/ccmo-notitie-mri-bij-kids-14-102004-eng.pdf. (Assessed, August 7, 2014)

Chandrasekaran, B., Kraus, N., & Wong, P. C. (2012). Human inferior colliculus activity relates to individual differences in spoken language learning. *Journal of Neurophysiology*, *107*(5), 1325.

Chang, S.-E., Erickson, K. I., Ambrose, N. G., Hasegawa-Johnson, M. A., & Ludlow, C. L. (2008). Brain anatomy differences in childhood stuttering. *Neuroimage*, *39*(3), 1333–1344.

Chew, S. J., Mello, C., Nottebohm, F., Jarvis, E., & Vicario, D. S. (1995). Decrements in auditory responses to a repeated conspecific song are long-lasting and require two periods of protein synthesis in the songbird forebrain. *Proceedings of the National Academy of Sciences*, *92*(8), 3406–3410.

Chomsky, N. (1957). *Syntactic structures*. Walter De Gruyter.

Clayton, N. (1988). Song discrimination learning in zebra finches. *Animal Behaviour*, *36*(4), 1016–1024.

Cowen, P. (1950). Strain differences in mice to the carcinogenic action of urethane and its non-carcinogenicity in chicks and guinea-pigs. *British Journal of Cancer*, *4*(2), 245.

Cowie, R., Douglas-Cowie, E., & Kerr, A. (1982). A study of speech deterioration in post-lingually deafened adults. *The Journal of Laryngology & Otology*, *96*(02), 101–112.

Curtiss, S., Fromkin, V., Krashen, S., Rigler, D., & Rigler, M. (1974). The linguistic development of genie. *Language*, 528–554.

Cynx, J., Williams, H., & Nottebohm, F. (1992). Hemispheric differences in avian song discrimination. *Proceedings of the National Academy of Sciences*, *89*(4), 1372–1375.

Dave, A. S., Albert, C. Y., & Margoliash, D. (1998). Behavioral state modulation of auditory activity in a vocal motor system. *Science*, *282*(5397), 2250–2254.

Dave, A. S., & Margoliash, D. (2000). Song replay during sleep and computational rules for sensorimotor vocal learning. *Science*, *290*(5492), 812–816.

DeCasper, A., & Spence, M. (1986). Newborns prefer a familiar story over an unfamiliar one. *Infant Behavior and Development*, *9*, 133–150.

De Groof, G., Jonckers, E., Güntürkün, O., Denolf, P., Van Auderkerke, J., & Van der Linden, A. (2013). Functional mri and functional connectivity of the visual system of awake pigeons. *Behavioural Brain Research*, *239*, 43–50.

De Groof, G., Poirier, C., George, I., Hausberger, M., & Van der Linden, A. (2011). Seasonal changes in selectivity for individual song in the songbird auditory forebrain. Abstract and Poster, SfN's 41st Annual Meeting, November 12-16, Washington, DC, USA.

De Groof, G., & Van der Linden, A. (2010). Love songs, bird brains and diffusion tensor imaging. *NMR in Biomedicine*, *23*(7), 873–883.

Dehaene-Lambertz, G., Dehaene, S., & Hertz-Pannier, L. (2002). Functional neuroimaging of speech perception in infants. *Science*, *298*(5600), 2013–2015.

Derégnaucourt, S., Poirier, C., Van der Kant, A., Van der Linden, A., & Gahr, M. (2013). Comparisons of different methods to train a young zebra finch (taeniopygia guttata) to learn a song. *Journal of Physiology-Paris*, *107*(3), 210–218.

De Vries, M. H., Petersson, K. M., Geukes, S., Zwitserlood, P., & Christiansen, M. H. (2012). Processing multiple non-adjacent dependencies: evidence from sequence learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1598), 2065–2076.

Doupe, A. J., & Konishi, M. (1991). Song-selective auditory circuits in the vocal control system of the zebra finch. *Proceedings of the National Academy of Sciences*, *88*(24), 11339–11343.

Doupe, A. J., & Kuhl, P. K. (1999). Birdsong and human speech: common themes and mechanisms. *Annual Review of Neuroscience*, *22*(1), 567–631.

Duncan, K. A., & Saldanha, C. J. (2011). Neuroinflammation induces glial aromatase expression in the uninjured songbird brain. *Journal of Neuroinflammation*, *8*(1), 1–11.

Duong, T. Q., Yacoub, E., Adriany, G., Hu, X., Uurbil, K., & Kim, S.-G. (2003). Microvascular bold contribution at 4 and 7 t in the human brain: Gradient-echo and spin-echo fmri with suppression of blood effects. *Magnetic Resonance in Medicine*, *49*(6), 1019–1027.

Dutka, M. V., Scanley, B. E., Does, M. D., & Gore, J. C. (2002). Changes in cbf-bold coupling detected by mri during and after repeated transient hypercapnia in rat. *Magnetic Resonance in Medicine*, *48*(2), 262–270.

Eales, L. A. (1985). Song learning in zebra finches: some effects of song model availability on what is learnt and when. *Animal Behaviour*, *33*(4), 1293–1300.

Eens, M., Pinxten, R., & Verheyen, R. F. (1992). Song learning in captive european starlings, sturnus vulgaris. *Animal Behaviour*, *44*(6), 1131–1143.

Egnor, S., & Hauser, M. D. (2004). A paradox in the evolution of primate vocal learning. *Trends in Neurosciences*, *27*(11), 649–654.

Eliades, S. J., & Wang, X. (2008). Neural substrates of vocalization feedback monitoring in primate auditory cortex. *Nature*, *453*(7198), 1102–1106.

Ellis, N. C. (1996). Working memory in the acquisition of vocabulary and syntax: Putting language in good order. *The Quarterly Journal of Experimental Psychology: Section A*, *49*(1), 234–250.

Ferris, C. F., Kulkarni, P., Sullivan, J. M., Harder, J. A., Messenger, T. L., & Febo, M. (2005). Pup suckling is more rewarding than cocaine: evidence from functional magnetic resonance imaging and three-dimensional computational analysis. *The Journal of Neuroscience*, *25*(1), 149–156.

Ferris, C. F., Stolberg, T., Kulkarni, P., Murugavel, M., Blanchard, R., Blanchard, D. C., . . . Simon, N. G. (2008). Imaging the neural circuitry and chemical control of aggressive motivation. *BMC Neuroscience*, *9*(1), 111.

Fields, R. D. (2005). Myelination: an overlooked mechanism of synaptic plasticity? *The Neuroscientist*, *11*(6), 528–531.

Fischer, J., & Hammerschmidt, K. (2011). Ultrasonic vocalizations in mouse models for speech and socio-cognitive disorders: insights into the evolution of vocal communication. *Genes, Brain and Behavior*, *10*(1), 17–27.

Flöel, A., De Vries, M. H., Scholz, J., Breitenstein, C., & Johansen-Berg, H. (2009). White matter integrity in the vicinity of broca's area predicts grammar learning success. *Neuroimage*, *47*(4), 1974–1981.

Folia, V., & Petersson, K. M. (2014). Implicit structured sequence learning: An fmri study of the structural mere-exposure effect. *Frontiers in Psychology*, *5*, 41.

Friederici, A. D., & Gierhan, S. M. (2013). The language network. *Current Opinion in Neurobiology*, *23*(2), 250–254.

Friederici, A. D., Rueschemeyer, S.-A., Hahne, A., & Fiebach, C. J. (2003). The role of left inferior frontal and superior temporal cortex in sentence comprehension: localizing syntactic and semantic processes. *Cerebral Cortex*, *13*(2), 170–177.

Fromkin, V., Krashen, S., Curtiss, S., Rigler, D., & Rigler, M. (1974). The development of language in genie: a case of language acquisition beyond the critical period. *Brain and Language*, *1*(1), 81–107.

Funabiki, Y., & Konishi, M. (2003). Long memory in song learning by zebra finches. *The Journal of Neuroscience*, *23*(17), 6928–6935.

Gentner, T. Q., Fenn, K. M., Margoliash, D., & Nusbaum, H. C. (2006). Recursive syntactic pattern learning by songbirds. *Nature*, *440*(7088), 1204–1207.

George, I., Cousillas, H., Richard, J.-P., & Hausberger, M. (2005). Auditory responses in the hvc of anesthetized starlings. *Brain Research*, *1047*(2), 245–260.

Gervain, J., Macagno, F., Cogoi, S., Peña, M., & Mehler, J. (2008). The neonate brain detects speech structure. *Proceedings of the National Academy of Sciences*, *105*(37), 14222–14227.

Geschwind, N. (1965). Disconnexion syndromes in animal and man. *Brain*, *88*, 237–294.

Gobes, S. M., & Bolhuis, J. J. (2007). Birdsong memory: a neural dissociation between song recognition and production. *Current Biology*, *17*(9), 789–793.

Gobes, S. M., Ter Haar, S. M., Vignal, C., Vergne, A. L., Mathevon, N., & Bolhuis, J. J. (2009). Differential responsiveness in brain and behavior to sexually dimorphic long calls in male and female zebra finches. *Journal of Comparative Neurology*, *516*(4), 312–320.

Gobes, S. M., Zandbergen, M. A., & Bolhuis, J. J. (2010). Memory in the mak-

ing: localized brain activation related to song learning in young song-birds. *Proceedings of the Royal Society B: Biological Sciences*, *277*(1698), 3343–3351.

Goldstein, M. H., King, A. P., & West, M. J. (2003). Social interaction shapes babbling: Testing parallels between birdsong and speech. *Proceedings of the National Academy of Sciences*, *100*(13), 8030–8035.

Gómez, R. (2002). Variability and detection of invariant structure. *Psychological Science*, *13*(5), 431–436.

Gómez, R., & Maye, J. (2005). The developmental trajectory of nonadjacent dependency learning. *Infancy*, *7*(2), 183–206.

Grace, J. A., Amin, N., Singh, N. C., & Theunissen, F. E. (2003). Selectivity for conspecific song in the zebra finch auditory forebrain. *Journal of Neurophysiology*, *89*(1), 472–487.

Grama, I., Wijnen, F., & Kerkhoff, A. (2013). Constraints on non-adjacent dependency-learning: Distance matters an artificial grammar learning study with adults. In S. Baiz, N. Goldman, & R. Hawkes (Eds.), *Boston university conference on language development 37 online proceedings supplement*. Boston, Massachussetts.

Greicius, M. D., Supekar, K., Menon, V., & Dougherty, R. F. (2009). Resting-state functional connectivity reflects structural connectivity in the default mode network. *Cerebral Cortex*, *19*(1), 72–78.

Güntürkün, O., Verhoye, M., De Groof, G., & Van der Linden, A. (2013). A 3-dimensional digital atlas of the ascending sensory and the descending motor systems in the pigeon brain. *Brain Structure and Function*, *218*(1), 269–281.

Haesler, S., Rochefort, C., Georgi, B., Licznerski, P., Osten, P., & Scharff, C. (2007). Incomplete and inaccurate vocal imitation after knockdown of foxp2 in songbird basal ganglia nucleus area x. *PLoS Biology*, *5*(12), e321.

Hakuta, K., Bialystok, E., & Wiley, E. (2003). Critical evidence a test of the critical-period hypothesis for second-language acquisition. *Psychological Science*, *14*(1), 31–38.

Hara, E., Rivas, M. V., Ward, J. M., Okanoya, K., & Jarvis, E. D. (2012). Convergent differential regulation of parvalbumin in the brains of vocal learners. *PloS One*, *7*(1), e29457.

Harel, N., Lin, J., Moeller, S., Ugurbil, K., & Yacoub, E. (2006). Combined imaging–histological study of cortical laminar specificity of fmri signals. *Neuroimage*, *29*(3), 879–887.

Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science*, *298*(5598), 1569–1579.

Hein, G., Doehrmann, O., Müller, N. G., Kaiser, J., Muckli, L., & Naumer, M. J. (2007). Object familiarity and semantic congruency modulate responses in cortical audiovisual integration areas. *The Journal of Neuroscience*, *27*(30), 7881–7887.

Hickok, G. (2009). The functional neuroanatomy of language. *Physics of Life Reviews*, *6*(3), 121–143.

Hickok, G., & Poeppel, D. (2000). Towards a functional neuroanatomy of speech perception. *Trends in cognitive sciences*, *4*(4), 131–138.

Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition*, *92*(1), 67–99.

Hilliard, A. T., Miller, J. E., Fraley, E. R., Horvath, S., & White, S. A. (2012). Molecular microcircuitry underlies functional specification in a basal ganglia circuit dedicated to vocal learning. *Neuron*, *73*(3), 537–552.

Hockett, C. F. (1963). The problem of universals in language. *Universals of Language*, *2*, 1–29.

Houde, J., Nagarajan, S., Sekihara, K., & Merzenich, M. (2002). Modulation of the auditory cortex during speech: an meg study. *Journal of Cognitive Neuroscience*, *14*(8), 1125–1138.

Hsu, A., Woolley, S. M., Fremouw, T. E., & Theunissen, F. E. (2004). Modulation power and phase spectrum of natural sounds enhance neural encoding performed by single auditory neurons. *The Journal of Neuroscience*, *24*(41), 9201–9211.

Iacoboni, M. (2005). Neural mechanisms of imitation. *Current Opinion in Neurobiology*, *15*(6), 632–637.

Janata, P., & Margoliash, D. (1999). Gradual emergence of song selectivity in sensorimotor structures of the male zebra finch song system. *The Journal of Neuroscience*, *19*(12), 5108–5118.

Jäncke, L., Shah, N., Posse, S., Grosse-Ryuken, M., & Müller-Gärtner, H.-W. (1998). Intensity coding of auditory stimuli: an fmri study. *Neuropsychologia*, *36*(9), 875–883.

Janik, V., & Slater, P. B. (1997). Vocal learning in mammals. *Advances in the Study of Behavior*, *26*, 59–99.

Jarvis, E. D. (2004). Learned birdsong and the neurobiology of human language. *Annals of the New York Academy of Sciences*, *1016*(1), 749–777.

Jarvis, E. D., Güntürkün, O., Bruce, L., Csillag, A., Karten, H., Kuenzel, W., . . . Butler, A. B. (2005). Avian brains and a new understanding of vertebrate brain evolution. *Nature Reviews Neuroscience*, *6*(2), 151–159.

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of english as a second language. *Cognitive Psychology*, *21*(1), 60–99.

Jonckers, E., Van Audekerke, J., De Visscher, G., Van der Linden, A., & Verhoye, M. (2011). Functional connectivity fmri of the rodent brain: comparison of functional connectivity networks in rat and mouse. *PloS One*, *6*(4), e18876.

Jones, M., Berwick, J., Hewson-Stoate, N., Gias, C., & Mayhew, J. (2005). The effect of hypercapnia on the neural and hemodynamic responses to somatosensory stimulation. *Neuroimage*, *27*(3), 609–623.

Kana, R. K., Keller, T. A., Cherkassky, V. L., Minshew, N. J., & Just, M. A. (2006). Sentence comprehension in autism: thinking in pictures with decreased functional connectivity. *Brain*, *129*(9), 2484–2493.

Kao, M. H., Doupe, A. J., & Brainard, M. S. (2005). Contributions of an avian basal ganglia–forebrain circuit to real-time modulation of song. *Nature*, *433*(7026), 638–643.

Keenan, J. P., Nelson, A., O'Connor, M., & Pascual-Leone, A. (2001). Self-recognition and the right hemisphere. *Nature*.

Keller, G. B., & Hahnloser, R. H. (2009). Neural processing of auditory feedback during vocal practice in a songbird. *Nature*, *457*(7226), 187–190.

Kerkhoff, A., De Bree, E., De Klerk, M., & Wijnen, F. (2013). Non-adjacent dependency learning in infants at familial risk of dyslexia. *Journal of Child Language*, *40*(01), 11–28.

Knörnschild, M., Behr, O., & Von Helversen, O. (2006). Babbling behavior in the sac-winged bat (saccopteryx bilineata). *Naturwissenschaften*, *93*(9), 451–454.

Konishi, M. (1965). The role of auditory feedback in the control of vocalization in the white-crowned sparrow1. *Zeitschrift für Tierpsychologie*, *22*(7), 770–783.

Kotz, S. A., Frisch, S., Von Cramon, D., & Friederici, A. D. (2003). Syntactic language processing: Erp lesion data on the role of the basal ganglia. *Journal of the International Neuropsychological Society*, *9*(07), 1053–1060.

Koyama, M. S., Di Martino, A., Zuo, X.-N., Kelly, C., Mennes, M., Jutagir, D. R., ... Milham, M. P. (2011). Resting-state functional connectivity indexes reading competence in children and adults. *The Journal of Neuroscience*, *31*(23), 8617–8624.

Kuhl, P. K. (2003). Human speech and birdsong: communication and the social brain. *Proceedings of the National Academy of Sciences*, *100*(17), 9645–9646.

Lai, C. S., Fisher, S. E., Hurst, J. A., Vargha-Khadem, F., & Monaco, A. P. (2001). A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature*, *413*(6855), 519–523.

Lancaster, J. L., Woldorff, M. G., Parsons, L. M., Liotti, M., Freitas, C. S., Rainey, L., ... Fox, P. T. (2000). Automated talairach atlas labels for functional brain mapping. *Human Brain Mapping*, *10*(3), 120–131.

Lauay, C., Gerlach, N. M., Adkins-Regan, E., & DeVoogd, T. J. (2004). Female zebra finches require early song exposure to prefer high-quality song as adults. *Animal Behaviour*, *68*(6), 1249–1255.

Le Bihan, D., Mangin, J.-F., Poupon, C., Clark, C. A., Pappata, S., Molko, N., & Chabriat, H. (2001). Diffusion tensor imaging: concepts and applications. *Journal of Magnetic Resonance Imaging*, *13*(4), 534–546.

Lee, S. P., Duong, T. Q., Yang, G., Iadecola, C., & Kim, S. G. (2001). Relative changes of cerebral arterial and venous blood volumes during increased cerebral blood flow: implications for bold fmri. *Magnetic Resonance in*

*Medicine*, *45*(5), 791–800.

Lee, S. P., Silva, A. C., Ugurbil, K., & Kim, S. G. (1999). Diffusion-weighted spin-echo fmri at 9.4 t: microvascular/tissue contribution to bold signal changes. *Magnetic Resonance in Medicine*, *42*(5), 919–928.

Leemans, A., Jeurissen, B., Sijbers, J., & Jones, D. (2009). Exploredti: a graphical toolbox for processing, analyzing, and visualizing diffusion mr data. In *17th annual meeting of intl soc mag reson med* (p. 3537).

Lei, H., & Mooney, R. (2010). Manipulation of a central auditory representation shapes learned vocal output. *Neuron*, *65*(1), 122–134.

Lenneberg, E. H., Chomsky, N., & Marx, O. (1967). *Biological foundations of language* (Vol. 68). Wiley New York.

Lenneberg, E. H., Rebelsky, F. G., & Nichols, I. A. (1965). The vocalizations of infants born to deaf and to hearing parents. *Human Development*, *8*(1), 23–37.

Leonardo, A., & Konishi, M. (1999). Decrystallization of adult birdsong by perturbation of auditory feedback. *Nature*, *399*(6735), 466–470.

Lewicki, M. S., & Arthur, B. J. (1996). Hierarchical organization of auditory temporal context sensitivity. *The Journal of Neuroscience*, *16*(21), 6987–6998.

Li, P., Legault, J., & Litcofsky, K. A. (2014). Neuroplasticity as a function of second language learning: Anatomical changes in the human brain. *Cortex*.

Liang, Z., King, J., & Zhang, N. (2011). Uncovering intrinsic connectional architecture of functional networks in awake rat brain. *The Journal of Neuroscience*, *31*(10), 3776–3783.

Lin, L., Chen, K., Alexander, G., He, J., Valla, J., Galons, J., . . . Reiman, E. (2003). Construction of mouse brain mri templates using spm99. In *Proceedings of the 5th ifac symposium on modelling and control in biomedical systems, melbourne, australia* (pp. 113–118).

Logothetis, N. K., Guggenberger, H., Peled, S., & Pauls, J. (1999). Functional imaging of the monkey brain. *Nature Neuroscience*, *2*(6), 555–562.

Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., & Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fmri signal. *Nature*, *412*(6843), 150–157.

Lohmann, G., Hoehl, S., Brauer, J., Danielmeier, C., Bornkessel-Schlesewsky, I., Bahlmann, J., . . . Friederici, A. (2010). Setting the frame: the human brain activates a basic low-frequency network for language processing. *Cerebral Cortex*, *20*(6), 1286–1292.

London, S. E., & Clayton, D. F. (2008). Functional identification of sensory mechanisms required for developmental song learning. *Nature Neuroscience*, *11*(5), 579–586.

Lopez-Barroso, D., de Diego-Balaguer, R., Cunillera, T., Camara, E., Münte, T. F., & Rodriguez-Fornells, A. (2011). Language learning under working memory constraints correlates with microstructural differences in

the ventral language pathway. *Cerebral Cortex*, *21*(12), 2742–2750.

Loui, P., Li, H. C., & Schlaug, G. (2011). White matter integrity in right hemisphere predicts pitch-related grammar learning. *Neuroimage*, *55*(2), 500–507.

MacKenzie-Graham, A., Lee, E. F., Dinov, I. D., Bota, M., Shattuck, D. W., Ruffins, S., . . . Toga, A. W. (2004). A multimodal, multidimensional atlas of the c57bl/6j mouse brain. *Journal of Anatomy*, *204*(2), 93–102.

Malonek, D., Dirnagl, U., Lindauer, U., Yamada, K., Kanno, I., & Grinvald, A. (1997). Vascular imprints of neuronal activity: relationships between the dynamics of cortical blood flow, oxygenation, and volume changes following sensory stimulation. *Proceedings of the National Academy of Sciences*, *94*(26), 14826–14831.

Mano, Y., Sugiura, M., Tsukiura, T., Chiao, J. Y., Yomogida, Y., Jeong, H., . . . Kawashima, R. (2011). The representation of social interaction in episodic memory: A functional mri study. *NeuroImage*, *57*(3), 1234–1242.

Marcus, G. F., Vijayan, S., Rao, S. B., & Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science*, *283*(5398), 77–80.

Margoliash, D. (1986). Preference for autogenous song by auditory neurons in a song system nucleus of the white-crowned sparrow. *The Journal of Neuroscience*, *6*(6), 1643–1661.

Margoliash, D., & Fortune, E. S. (1992). Temporal and harmonic combination-sensitive neurons in the zebra finch's hvc. *The Journal of Neuroscience*, *12*(11), 4309–4326.

Margoliash, D., & Schmidt, M. F. (2010). Sleep, off-line processing, and vocal learning. *Brain and Language*, *115*(1), 45–58.

Marler, P. (1970). Birdsong and speech development: could there be parallels? *American Scientist*, *58*, 669–673.

Martin, C., Jones, M., Martindale, J., & Mayhew, J. (2006). Haemodynamic and neural responses to hypercapnia in the awake rat. *European Journal of Neuroscience*, *24*(9), 2601–2610.

Masamoto, K., Kim, T., Fukuda, M., Wang, P., & Kim, S.-G. (2007). Relationship between neural, vascular, and bold signals in isoflurane-anesthetized rat somatosensory cortex. *Cerebral Cortex*, *17*(4), 942–950.

Matsuo, K., Chen, S. H. A., & Tseng, W. Y. I. (2012). Aveli: A robust lateralization index in functional magnetic resonance imaging using unbiased threshold-free computation. *Journal of Neuroscience Methods*, *205*(1), 119–129.

Maul, K. K., Voss, H. U., Parra, L. C., Salgado-Commissariat, D., Ballon, D., Tchernichovski, O., & Helekar, S. A. (2010). The development of stimulus-specific auditory responses requires song exposure in male but not female zebra finches. *Developmental Neurobiology*, *70*(1), 28–40.

Mazziotta, J., Toga, A., Evans, A., Fox, P., Lancaster, J., Zilles, K., . . . Mazoyer, B. (2001). A probabilistic atlas and reference system for the human brain: International consortium for brain mapping (icbm). *Philosophical Trans-*

*actions of the Royal Society of London. Series B: Biological Sciences*, *356*(1412), 1293–1322.

McNealy, K., Mazziotta, J. C., & Dapretto, M. (2006). Cracking the language code: Neural mechanisms underlying speech parsing. *The Journal of Neuroscience*, *26*(29), 7629–7639.

McNealy, K., Mazziotta, J. C., & Dapretto, M. (2010). The neural basis of speech parsing in children and adults. *Developmental Science*, *13*(2), 385–406.

Mello, C. V. (2004). Identification and analysis of vocal communication pathways in birds through inducible gene expression. *Anais da Academia Brasileira de Ciências*, *76*(2), 243–246.

Mello, C. V., Vicario, D. S., & Clayton, D. F. (1992). Song presentation induces gene expression in the songbird forebrain. *Proceedings of the National Academy of Sciences*, *89*(15), 6818–6822.

Meltzer, J. A., Postman-Caucheteux, W. A., McArdle, J. J., & Braun, A. R. (2009). Strategies for longitudinal neuroimaging studies of overt language production. *Neuroimage*, *47*(2), 745–755.

Minagawa-Kawai, Y., Cristià, A., & Dupoux, E. (2011). Cerebral lateralization and early speech acquisition: a developmental scenario. *Developmental Cognitive Neuroscience*, *1*(3), 217–232.

Mirzatoni, A., Spence, R. D., Naranjo, K. C., Saldanha, C. J., & Schlinger, B. A. (2010). Injury-induced regulation of steroidogenic gene expression in the cerebellum. *Journal of Neurotrauma*, *27*(10), 1875–1882.

Moelker, A., & Pattynama, P. M. (2003). Acoustic noise concerns in functional magnetic resonance imaging. *Human Brain Mapping*, *20*(3), 123–141.

Moon, C., Cooper, R. P., & Fifer, W. P. (1993). Two-day-olds prefer their native language. *Infant Behavior and Development*, *16*(4), 495–500.

Moorman, S., Gobes, S. M., Kuijpers, M., Kerkhofs, A., Zandbergen, M. A., & Bolhuis, J. J. (2012). Human-like brain hemispheric dominance in birdsong learning. *Proceedings of the National Academy of Sciences*, *109*(31), 12782–12787.

Moseley, M. (2002). Diffusion tensor imaging and aging–a review. *NMR in Biomedicine*, *15*(7-8), 553–560.

Mower, G. D. (1991). The effect of dark rearing on the time course of the critical period in cat visual cortex. *Developmental Brain Research*, *58*(2), 151–158.

Müller-Preuss, P., & Ploog, D. (1981). Inhibition of auditory cortical neurons during phonation. *Brain Research*, *215*(1), 61–76.

Nakamura, K., Kawashima, R., Sugiura, M., Kato, T., Nakamura, A., Hatano, K., . . . Kojimaa, S. (2001). Neural substrates for recognition of familiar voices: a pet study. *Neuropsychologia*, *39*(10), 1047–1054.

Newport, E. L., & Aslin, R. N. (2004). Learning at a distance i. statistical learning of non-adjacent dependencies. *Cognitive Psychology*, *48*(2), 127–162.

Nieuwenhuis, I. L., Folia, V., Forkstam, C., Jensen, O., & Petersson, K. M.

(2013). Sleep promotes the extraction of grammatical rules. *PloS One*, *8*(6), e65046.

Nikolopoulos, T. P., O'Donoghue, G. M., & Archbold, S. (1999). Age at implantation: its importance in pediatric cochlear implantation. *The Laryngoscope*, *109*(4), 595–599.

Nilson, P. C., Teramitsu, I., & White, S. A. (2005). Caudal thoracic air sac cannulation in zebra finches for isoflurane anesthesia. *Journal of Neuroscience Methods*, *143*(2), 107–115.

Nordeen, K. W., & Nordeen, E. J. (1992). Auditory feedback is necessary for the maintenance of stereotyped song in adult zebra finches. *Behavioral and Neural Biology*, *57*(1), 58–66.

Nordeen, K. W., & Nordeen, E. J. (1993). Long-term maintenance of song in adult zebra finches is not affected by lesions of a forebrain region involved in song learning. *Behavioral and Neural Biology*, *59*(1), 79–82.

Nottebohm, F., & Nottebohm, M. E. (1978). Relationship between song repertoire and age in the canary, serinus canarius. *Zeitschrift für Tierpsychologie*, *46*(3), 298–305.

Numminen, J., Salmelin, R., & Hari, R. (1999). Subject's own speech reduces reactivity of the human auditory cortex. *Neuroscience Letters*, *265*(2), 119–122.

Obleser, J., Wise, R. J., Dresner, M. A., & Scott, S. K. (2007). Functional integration across brain regions improves speech perception under adverse listening conditions. *The Journal of Neuroscience*, *27*(9), 2283–2289.

Ogawa, S., Lee, T., Kay, A., & Tank, D. (1990). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences*, *87*(24), 9868–9872.

Ogawa, S., Tank, D. W., Menon, R., Ellermann, J. M., Kim, S. G., Merkle, H., & Ugurbil, K. (1992). Intrinsic signal changes accompanying sensory stimulation: functional brain mapping with magnetic resonance imaging. *Proceedings of the National Academy of Sciences*, *89*(13), 5951–5955.

Ölveczky, B. P., Andalman, A. S., & Fee, M. S. (2005). Vocal experimentation in the juvenile songbird requires a basal ganglia circuit. *PLoS Biology*, *3*(5), e153.

Onnis, L., Monaghan, P., Richmond, K., & Chater, N. (2005). Phonology impacts segmentation in online speech processing. *Journal of Memory and Language*, *53*(2), 225–237.

Opitz, B., & Friederici, A. D. (2003). Interactions of the hippocampal system and the prefrontal cortex in learning language-like rules. *NeuroImage*, *19*(4), 1730–1737.

Pacton, S., & Perruchet, P. (2008). An attention-based associative account of adjacent and nonadjacent dependency learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(1), 80.

Parker, G. J., Luzzi, S., Alexander, D. C., Wheeler-Kingshott, C. A., Ciccarelli, O., & Lambon Ralph, M. A. (2005). Lateralization of ventral and dor-

sal auditory-language pathways in the human brain. *Neuroimage*, *24*(3), 656–666.

Peeters, R., Tindemans, I., De Schutter, E., & Van der Linden, A. (2001). Comparing bold fmri signal changes in the awake and anesthetized rat during electrical forepaw stimulation. *Magnetic Resonance Imaging*, *19*(6), 821–826.

Peña, M., Bonatti, L. L., Nespor, M., & Mehler, J. (2002). Signal-driven computations in speech processing. *Science*, *298*(5593), 604–607.

Pepperberg, I. M. (2002). In search of king solomons ring: cognitive and communicative studies of grey parrots (psittacus erithacus). *Brain, Behavior and Evolution*, *59*(1-2), 54–67.

Perruchet, P., Tyler, M. D., Galland, N., & Peereman, R. (2004). Learning nonadjacent dependencies: no need for algebraic-like computations. *Journal of Experimental Psychology: General*, *133*(4), 573.

Peterson, R. S., Fernando, G., Day, L., Allen, T. A., Chapleau, J. D., Menjivar, J., . . . Lee, D. W. (2007). Aromatase expression and cell proliferation following injury of the adult zebra finch hippocampus. *Developmental Neurobiology*, *67*(14), 1867–1878.

Petersson, K.-M., Folia, V., & Hagoort, P. (2012). What artificial grammar learning reveals about the neurobiology of syntax. *Brain and Language*, *120*(2), 83–95.

Petersson, K. M., Forkstam, C., & Ingvar, M. (2004). Artificial syntactic violations activate brocas region. *Cognitive Science*, *28*(3), 383–407.

Petersson, K. M., & Hagoort, P. (2012). The neurobiology of syntax: beyond string sets. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1598), 1971–1983.

Petkov, C. I., & Jarvis, E. D. (2012). Birds, primates, and spoken language origins: behavioral phenotypes and neurobiological substrates. *Frontiers in Evolutionary Neuroscience*, *4*.

Petkov, C. I., Kayser, C., Augath, M., & Logothetis, N. K. (2006). Functional imaging reveals numerous fields in the monkey auditory cortex. *PLoS Biology*, *4*(7), e215.

Petkov, C. I., Kayser, C., Steudel, T., Whittingstall, K., Augath, M., & Logothetis, N. K. (2008). A voice region in the monkey brain. *Nature Neuroscience*, *11*(3), 367–374.

Pfenning, A. R., Hara, E., Whitney, O., Rivas, M. V., Wang, R., Roulhac, P. L., . . . others (2014). Convergent transcriptional specializations in the brains of humans and song-learning birds. *Science*, *346*(6215), 1256846.

Phan, M. L., Pytte, C. L., & Vicario, D. S. (2006). Early auditory experience generates long-lasting memories that may subserve vocal learning in songbirds. *Proceedings of the National Academy of Sciences*, *103*(4), 1088–1093.

Phan, M. L., & Vicario, D. S. (2010). Hemispheric differences in processing of vocalizations depend on early experience. *Proceedings of the National Academy of Sciences*, *107*(5), 2301–2306.

Pinaud, R., Fortes, A. F., Lovell, P., & Mello, C. V. (2006). Calbindin-positive neurons reveal a sexual dimorphism within the songbird analogue of the mammalian auditory cortex. *Journal of Neurobiology*, *66*(2), 182–195.

Poirier, C., Boumans, T., Vellema, M., De Groof, G., Charlier, T. D., Verhoye, M., ... Balthazart, J. (2011). Own song selectivity in the songbird auditory pathway: suppression by norepinephrine. *PloS One*, *6*(5), e20131.

Poirier, C., Boumans, T., Verhoye, M., Balthazart, J., & Van der Linden, A. (2009). Own-song recognition in the songbird auditory pathway: selectivity and lateralization. *The Journal of Neuroscience*, *29*(7), 2252–2258.

Poirier, C., Vellema, M., Verhoye, M., Van Meir, V., Wild, J. M., Balthazart, J., & Van der Linden, A. (2008). A three-dimensional mri atlas of the zebra finch brain in stereotaxic coordinates. *Neuroimage*, *41*(1), 1–6.

Poirier, C., Verhoye, M., Boumans, T., & Van der Linden, A. (2010). Implementation of spin-echo blood oxygen level-dependent (bold) functional mri in birds. *NMR in Biomedicine*, *23*(9), 1027–1032.

Pollick, A. S., & De Waal, F. B. (2007). Ape gestures and language evolution. *Proceedings of the National Academy of Sciences*, *104*(19), 8184–8189.

Poole, J. H., Tyack, P. L., Stoeger-Horwath, A. S., & Watwood, S. (2005). Animal behaviour: elephants are capable of vocal learning. *Nature*, *434*(7032), 455–456.

Portas, C. M., Krakow, K., Allen, P., Josephs, O., Armony, J. L., & Frith, C. D. (2000). Auditory processing across the sleep-wake cycle: simultaneous eeg and fmri monitoring in humans. *Neuron*, *28*(3), 991–999.

Prather, J. F., & Mooney, R. (2004). Neural correlates of learned song in the avian forebrain: simultaneous representation of self and others. *Current Opinion in Neurobiology*, *14*(4), 496–502.

Prather, J. F., Peters, S., Nowicki, S., & Mooney, R. (2008). Precise auditory–vocal mirroring in neurons for learned vocal communication. *Nature*, *451*(7176), 305–310.

Price, C. J. (2010). The anatomy of language: a review of 100 fmri studies published in 2009. *Annals of the New York Academy of Sciences*, *1191*(1), 62–88.

Price, P. H. (1979). Developmental determinants of structure in zebra finch song. *Journal of Comparative and Physiological Psychology*, *93*(2), 260.

Pytte, C. L., Parent, C., Wildstein, S., Varghese, C., & Oberlander, S. (2010). Deafening decreases neuronal incorporation in the zebra finch caudomedial nidopallium (ncm). *Behavioural Brain Research*, *211*(2), 141–147.

Ralls, K., Fiorelli, P., & Gish, S. (1985). Vocalizations and vocal mimicry in captive harbor seals, phoca vitulina. *Canadian Journal of Zoology*, *63*(5), 1050–1056.

Ranasinghe, K., Vrana, W., Matney, C., & Kilgard, M. (2013). Increasing diversity of neural responses to speech sounds across the central auditory pathway. *Neuroscience*, *252*, 80–97.

Rauschecker, J. P., & Scott, S. K. (2009). Maps and streams in the auditory

cortex: nonhuman primates illuminate human speech processing. *Nature Neuroscience*, *12*(6), 718–724.

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, *6*(6), 855–863.

Reiss, D., & McCowan, B. (1993). Spontaneous vocal mimicry and production by bottlenose dolphins (tursiops truncatus): Evidence for vocal learning. *Journal of Comparative Psychology*, *107*(3), 301.

Ribeiro, S., Loula, A., De Araújo, I., Gudwin, R., & Queiroz, J. (2007). Symbols are not uniquely human. *Biosystems*, *90*(1), 263–272.

Richards, C. (2002). Anaesthetic modulation of synaptic transmission in the mammalian cns. *British Journal of Anaesthesia*, *89*(1), 79–90.

Riebel, K. (2000). Early exposure leads to repeatable preferences for male song in female zebra finches. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *267*(1461), 2553–2558.

Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, *3*(2), 131–141.

Roberts, T. F., Gobes, S. M., Murugan, M., Ölveczky, B. P., & Mooney, R. (2012). Motor circuits are required to encode a sensory model for imitative learning. *Nature Neuroscience*, *15*(10), 1454–1459.

Rondinoni, C., Amaro Jr, E., Cendes, F., Dos Santos, A., & Salmon, C. (2013). Effect of scanner acoustic background noise on strict resting-state fmri. *Brazilian Journal of Medical and Biological Research*, *46*(4), 359–367.

Sachdev, R. N., Champney, G. C., Lee, H., Price, R. R., Pickens III, D. R., Morgan, V. L., ... Ebner, F. F. (2003). Experimental model for functional magnetic resonance imaging of somatic sensory cortex in the unanesthetized rat. *Neuroimage*, *19*(3), 742–750.

Saffran, J. R. (2001). The use of predictive dependencies in language learning. *Journal of Memory and Language*, *44*(4), 493–515.

Saffran, J. R. (2002). Constraints on statistical language learning. *Journal of Memory and Language*, *47*(1), 172–196.

Saleem, K. S., & Logothetis, N. K. (2007). *A combined mri and histology atlas of the rhesus monkey brain in stereotaxic coordinates*. Academic Press.

Sandoval, M., & Gómez, R. L. (2013). The development of nonadjacent dependency learning in natural and artificial languages. *Wiley Interdisciplinary Reviews: Cognitive Science*, *4*(5), 511–522.

Santelmann, L. M., & Jusczyk, P. W. (1998). Sensitivity to discontinuous dependencies in language learners: Evidence for limitations in processing space. *Cognition*, *69*(2), 105–134.

Santi, A., & Grodzinsky, Y. (2007). Working memory and syntax interact in broca's area. *Neuroimage*, *37*(1), 8–17.

Sanvito, S., Galimberti, F., & Miller, E. H. (2007). Observational evidences of vocal learning in southern elephant seals: a longitudinal study. *Ethology*, *113*(2), 137–146.

Saur, D., Kreher, B. W., Schnell, S., Kümmerer, D., Kellmeyer, P., Vry, M.-S., ...

Weiller, C. (2008). Ventral and dorsal pathways for language. *Proceedings of the National Academy of Sciences*, *105*(46), 18035–18040.

Scharff, C., & Nottebohm, F. (1991). A comparative study of the behavioral deficits following lesions of various parts of the zebra finch song system: implications for vocal learning. *The Journal of Neuroscience*, *11*(9), 2896–2913.

Schmidt, M. F. (2010). An iacuc perspective on songbirds and their use in neurobiological research. *ILAR Journal*, *51*(4), 424.

Schmidt, M. F., & Konishi, M. (1998). Gating of auditory responses in the vocal control system of awake songbirds. *Nature Neuroscience*, *1*(6), 513–518.

Schmithorst, V. J., Wilke, M., Dardzinski, B. J., & Holland, S. K. (2002). Correlation of white matter diffusivity and anisotropy with age during childhood and adolescence: A cross-sectional diffusion-tensor mr imaging study 1. *Radiology*, *222*(1), 212–218.

Schneider, D. M., & Woolley, S. M. (2010). Discrimination of communication vocalizations by single neurons and groups of neurons in the auditory midbrain. *Journal of Neurophysiology*, *103*(6), 3248–3265.

Schölvinck, M. L., Maier, A., Frank, Q. Y., Duyn, J. H., & Leopold, D. A. (2010). Neural basis of global resting-state fmri activity. *Proceedings of the National Academy of Sciences*, *107*(22), 10238–10243.

Scholz, J., Klein, M. C., Behrens, T. E., & Johansen-Berg, H. (2009). Training induces changes in white-matter architecture. *Nature Neuroscience*, *12*(11), 1370–1371.

Schumacher, J. W., Schneider, D. M., & Woolley, S. M. (2011). Anesthetic state modulates excitability but not spectral tuning or neural discrimination in single auditory midbrain neurons. *Journal of Neurophysiology*, *106*(2), 500–514.

Schwarz, A. J., Danckaert, A., Reese, T., Gozzi, A., Paxinos, G., Watson, C., . . . Bifone, A. (2006). A stereotaxic mri template set for the rat brain with tissue class distribution maps and co-registered anatomical atlas: application to pharmacological mri. *Neuroimage*, *32*(2), 538–550.

Schwarzbauer, C., Mildner, T., Heinke, W., Brett, M., & Deichmann, R. (2010). Dual echo epi–the method of choice for fmri in the presence of magnetic field inhomogeneities? *Neuroimage*, *49*(1), 316–326.

Schweinhardt, P., Fransson, P., Olson, L., Spenger, C., & Andersson, J. L. (2003). A template for spatial normalisation of mr images of the rat brain. *Journal of Neuroscience Methods*, *129*(2), 105–113.

Seyfarth, R. M., Cheney, D. L., & Marler, P. (1980). Monkey responses to three different alarm calls: evidence of predator classification and semantic communication. *Science*, *210*(4471), 801–803.

Shaevitz, S. S., & Theunissen, F. E. (2007). Functional connectivity between auditory areas field l and clm and song system nucleus hvc in anesthetized zebra finches. *Journal of Neurophysiology*, *98*(5), 2747–2764.

Shank, S. S., & Margoliash, D. (2009). Sleep and sensorimotor integration

during early vocal learning in a songbird. *Nature*, *458*(7234), 73–77.

Slater, P. J., Eales, L. A., & Clayton, N. (1988). Song learning in zebra finches (ilzeniopygia guttata): Progress and prospects. *Advances in the Study of Behavior*, *18*, 1.

Sohrabji, F., Nordeen, E. J., & Nordeen, K. W. (1990). Selective impairment of song learning following lesions of a forebrain nucleus in the juvenile zebra finch. *Behavioral and Neural Biology*, *53*(1), 51–63.

Solis, M. M., & Doupe, A. J. (2000). Compromised neural selectivity for song in birds with impaired sensorimotor learning. *Neuron*, *25*(1), 109–121.

Song, J. H., Skoe, E., Banai, K., & Kraus, N. (2012). Training to improve hearing speech in noise: biological mechanisms. *Cerebral Cortex*, *22*(5), 1180–1190.

Song, J. H., Skoe, E., Wong, P. C., & Kraus, N. (2008). Plasticity in the adult human auditory brainstem following short-term linguistic training. *Journal of Cognitive Neuroscience*, *20*(10), 1892–1902.

Sossinka, R., & Böhner, J. (1980). Song types in the zebra finch poephila guttata castanotis1. *Zeitschrift für Tierpsychologie*, *53*(2), 123–132.

Spierings, M. J., & Ten Cate, C. (2014). Zebra finches are sensitive to prosodic features of human speech. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1787), 20140480.

Stevens, A. A., & Weaver, K. E. (2009). Functional characteristics of auditory cortex in the blind. *Behavioural Brain Research*, *196*(1), 134–138.

Stobbe, N., Westphal-Fitch, G., Aust, U., & Fitch, W. T. (2012). Visual artificial grammar learning: comparative research on humans, kea (nestor notabilis) and pigeons (columba livia). *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1598), 1995–2006.

Stoel-Gammon, C., & Otomo, K. (1986). Babbling development of hearing-impaired and normally hearing subjects. *Journal of Speech and Hearing Disorders*, *51*(1), 33–41.

Stripling, R., Volman, S. F., & Clayton, D. F. (1997). Response modulation in the zebra finch neostriatum: relationship to nuclear gene regulation. *The Journal of Neuroscience*, *17*(10), 3883–3893.

Tchernichovski, O., Mitra, P. P., Lints, T., & Nottebohm, F. (2001). Dynamics of the vocal imitation process: how a zebra finch learns its song. *Science*, *291*(5513), 2564–2569.

Tchernichovski, O., Nottebohm, F., Ho, C. E., Pesaran, B., & Mitra, P. P. (2000). A procedure for an automated measurement of song similarity. *Animal Behaviour*, *59*(6), 1167–1176.

Ten Cate, C. (1991). Behaviour-contingent exposure to taped song and zebra finch song learning. *Animal Behaviour*, *42*(5), 857–859.

Ten Cate, C., & Okanoya, K. (2012). Revisiting the syntactic abilities of non-human animals: natural vocalizations and artificial grammar learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1598), 1984–1994.

Teramitsu, I., Poopatanapong, A., Torrisi, S., & White, S. A. (2010). Striatal foxp2 is actively regulated during songbird sensorimotor learning. *PloS One*, *5*(1), e8548.

Ter Haar, S. M. (2013). *Birds and babies: a comparison of the early development in vocal learners*. PhD thesis, Institute of Biology Leiden (IBL), Leiden University Centre for Linguistics (LUCL), Leiden Institute for Brain and Cognition (LIBC), Faculty of Science, Leiden University.

Terpstra, N. J., Bolhuis, J. J., & Den Boer-Visser, A. M. (2004). An analysis of the neural representation of birdsong memory. *The Journal of Neuroscience*, *24*(21), 4971–4977.

Tettamanti, M., Alkadhi, H., Moro, A., Perani, D., Kollias, S., & Weniger, D. (2002). Neural correlates for the acquisition of natural language syntax. *Neuroimage*, *17*(2), 700–709.

Theunissen, F. E., Amin, N., Shaevitz, S. S., Woolley, S., Fremouw, T., & Hauber, M. E. (2004). Song selectivity in the song system and in the auditory forebrain. *Annals of the New York Academy of Sciences*, *1016*(1), 222–245.

Theunissen, F. E., & Doupe, A. J. (1998). Temporal and spectral sensitivity of complex auditory neurons in the nucleus hvc of male zebra finches. *The Journal of Neuroscience*, *18*(10), 3786–3802.

Thode, C., Güttinger, H., & Darlison, M. G. (2008). Expression of the gaba(a) receptor gamma4-subunit gene in discrete nuclei within the zebra finch song system. *Neuroscience*, *157*(1), 143–152.

Thorpe, W. H. (1958). The learning of song patterns by birds, with especial reference to the song of the chaffinch fringilla coelebs. *Ibis*, *100*(4), 535–570.

Tomasello, M., & Farrar, M. J. (1986). Joint attention and early language. *Child Development*, 1454–1463.

Tomasi, D., & Volkow, N. D. (2012). Resting functional connectivity of language networks: characterization and reproducibility. *Molecular Psychiatry*, *17*(8), 841–854.

Turk-Browne, N. B., Scholl, B. J., Chun, M. M., & Johnson, M. K. (2009). Neural evidence of statistical learning: Efficient detection of visual regularities without awareness. *Journal of Cognitive Neuroscience*, *21*(10), 1934–1945.

Tzounopoulos, T., & Kraus, N. (2009). Learning to encode timing: mechanisms of plasticity in the auditory brainstem. *Neuron*, *62*(4), 463–469.

Uddén, J., & Bahlmann, J. (2012). A rostro-caudal gradient of structured sequence processing in the left inferior frontal gyrus. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1598), 2023–2032.

Uddén, J., Folia, V., Forkstam, C., Ingvar, M., Fernández, G., Overeem, S., . . . Petersson, K. M. (2008). The inferior frontal cortex in artificial syntax processing: An rtms study. *Brain Research*, *1224*, 69–78.

Uddén, J., Ingvar, M., Hagoort, P., & Petersson, K. M. (2012). Implicit acquisition of grammars with crossed and nested non-adjacent dependencies:

Investigating the push-down stack model. *Cognitive Science*, *36*(6), 1078–1101.

Uludağ, K., Müller-Bierl, B., & Uğurbil, K. (2009). An integrative model for neuronal activity-induced signal changes for gradient and spin echo functional imaging. *Neuroimage*, *48*(1), 150–165.

Vallée, E., Commowick, O., Maumet, C., Stamm, A., Le Rumeur, E., Allaire, C., . . . Barillot, C. (2014). Statistical analysis of white matter integrity for the clinical study of typical specific language impairment in children. In T. Schultz, G. Nedjati-Gilani, A. Venkataraman, L. O'Donnell, & E. Panagiotaki (Eds.), *Computational diffusion mri and brain connectivity* (pp. 187–195). Springer.

Van den Bos, E., Christiansen, M. H., & Misyak, J. B. (2012). Statistical learning of probabilistic nonadjacent dependencies by multiple-cue integration. *Journal of Memory and Language*, *67*(4), 507–520.

Van der Kant, A., Derégnaucourt, S., Gahr, M., Van der Linden, A., & Poirier, C. (2013). Representation of early sensory experience in the adult auditory midbrain: Implications for vocal learning. *PloS One*, *8*(4), e61764.

Van der Kant, A., & Van der Linden, A. (2012). Neural correlates of song perception during zebra finch song learning. Abstract and Poster, 9th International Conference on the Evolution of Language (Evolang IX) , March 13-16, Kyoto, Japan.

Van der Linden, A., Van Camp, N., Ramos-Cabrer, P., & Hoehn, M. (2007). Current status of functional mri on small animals: application to physiology, pathophysiology, and cognition. *NMR in Biomedicine*, *20*(5), 522–545.

Van der Linden, A., Van Meir, V., Boumans, T., Poirier, C., & Balthazart, J. (2009). Mri in small brains displaying extensive plasticity. *Trends in Neurosciences*, *32*(5), 257–266.

Van der Linden, A., Verhoye, M., Van Audekerke, J., Peeters, R., Eens, M., Newman, S., . . . DeVoogd, T. (1998). Non invasive in vivo anatomical studies of the oscine brain by high resolution mri microscopy. *Journal of Neuroscience Methods*, *81*(1), 45–52.

Van Heijningen, C. A., De Visser, J., Zuidema, W., & Ten Cate, C. (2009). Simple rules can explain discrimination of putative recursive syntactic structures by a songbird species. *Proceedings of the National Academy of Sciences*, *106*(48), 20538–20543.

Van Heugten, M., & Johnson, E. K. (2010). Linking infants distributional learning abilities to natural language acquisition. *Journal of Memory and Language*, *63*(2), 197–209.

Vanhoutte, G., Verhoye, M., & Van der Linden, A. (2006). Changing body temperature affects the t 2* signal in the rat brain and reveals hypothalamic activity. *Magnetic Resonance in Medicine*, *55*(5), 1006–1012.

Van Meir, V., Boumans, T., De Groof, G., Van Audekerke, J., Smolders, A., Scheunders, P., . . . Van der Linden, A. (2005). Spatiotemporal properties of the bold response in the songbirds' auditory circuit during a variety of

listening tasks. *Neuroimage*, *25*(4), 1242–1255.

Van Meir, V., Verhoye, M., Absil, P., Eens, M., Balthazart, J., & Van der Linden, A. (2004). Differential effects of testosterone on neuronal populations and their connections in a sensorimotor brain nucleus controlling song production in songbirds: a manganese enhanced-magnetic resonance imaging study. *Neuroimage*, *21*(3), 914–923.

Van Ruijssevelt, L., De Groof, G., Van der Kant, A., Poirier, C., Van Audekerke, J., Verhoye, M., & Van der Linden, A. (2012). Functional magnetic resonance imaging (fmri) with auditory stimulation in songbirds. *Journal of Visualized Experiments*(76).

Vates, G. E., Broome, B. M., Mello, C. V., & Nottebohm, F. (1996). Auditory pathways of caudal telencephalon and their relation to the song system of adult male zebra finches (taenopygia guttata). *Journal of Comparative Neurology*, *366*(4), 613–642.

Vellema, M., Verschueren, J., Van Meir, V., & Van der Linden, A. (2011). A customizable 3-dimensional digital atlas of the canary brain in multiple modalities. *Neuroimage*, *57*(2), 352–361.

Verhoye, M., Van der Linden, A., Van Audekerke, J., Sijbers, J., Eens, M., & Balthazart, J. (1998). Imaging birds in a bird cage: in-vivo fse 3d mri of bird brain. *Magnetic Resonance Materials in Physics, Biology and Medicine*, *6*(1), 22–27.

Veroude, K., Norris, D. G., Shumskaya, E., Gullberg, M., & Indefrey, P. (2010). Functional connectivity between brain regions involved in learning words of a new language. *Brain and Language*, *113*(1), 21–27.

Vignal, C., Attia, J., Mathevon, N., & Beauchaud, M. (2004). Background noise does not modify song-induced genic activation in the bird brain. *Behavioural Brain Research*, *153*(1), 241–248.

Vignal, C., Boumans, T., Montcel, B., Ramstein, S., Verhoye, M., Van Audekerke, J., . . . Mottin, S. (2008). Measuring brain hemodynamic changes in a songbird: responses to hypercapnia measured with functional mri and near-infrared spectroscopy. *Physics in Medicine and Biology*, *53*(10), 2457.

Voss, H. U. (2011). A new technique for functional imaging in songbirds and beyond. *Journal of Cerebral Blood Flow & Metabolism*, *31*(2), 391.

Voss, H. U., Salgado-Commissariat, D., & Helekar, S. A. (2010). Altered auditory bold response to conspecific birdsong in zebra finches with stuttered syllables. *PloS One*, *5*(12), e14415.

Voss, H. U., Tabelow, K., Polzehl, J., Tchernichovski, O., Maul, K. K., Salgado-Commissariat, D., . . . Helekar, S. A. (2007). Functional mri of the zebra finch brain during song stimulation suggests a lateralized response topography. *Proceedings of the National Academy of Sciences*, *104*(25), 10667–10672.

Walters, B. J., Alexiades, N. G., & Saldanha, C. J. (2011). Intracerebral estrogen provision increases cytogenesis and neurogenesis in the injured zebra finch brain. *Developmental Neurobiology*, *71*(2), 170–181.

Warren, W. C., Clayton, D. F., Ellegren, H., Arnold, A. P., Hillier, L. W., Künstner, A., ... Wilson, R. K. (2010). The genome of a songbird. *Nature*, *464*(7289), 757–762.

Wartenburger, I., Heekeren, H. R., Abutalebi, J., Cappa, S. F., Villringer, A., & Perani, D. (2003). Early setting of grammatical processing in the bilingual brain. *Neuron*, *37*(1), 159–170.

Waterman, S. A., & Harding, C. F. (2008). Neurotoxic effects of dsp-4 on the central noradrenergic system in male zebra finches. *Behavioural Brain Research*, *188*(2), 271–280.

Werker, J. F., & Tees, R. C. (1983). Developmental changes across childhood in the perception of non-native speech sounds. *Canadian Journal of Psychology*, *37*(2), 278.

Westra, A. E., Zegers, M. P., Sukhai, R. N., Kaptein, A. A., Holscher, H. C., Ballieux, B. E., ... Wit, J. M. (2011). Discomfort in children undergoing unsedated mri. *European Journal of Pediatrics*, *170*(6), 771–777.

White, E. J., Genesee, F., & Steinhauer, K. (2012). Brain responses before and after intensive second language learning: proficiency based changes and first language background effects in adult learners. *PloS One*, *7*(12), e52318.

Whitfield-Gabrieli, S., & Nieto-Castanon, A. (2012). Conn: a functional connectivity toolbox for correlated and anticorrelated brain networks. *Brain Connectivity*, *2*(3), 125–141.

Wilbrecht, L., & Nottebohm, F. (2003). Vocal learning in birds and humans. *Mental Retardation and Developmental Disabilities Research Reviews*, *9*(3), 135–148.

Wild, J. (2004). Functional neuroanatomy of the sensorimotor control of singing. *Annals of the New York Academy of Sciences*, *1016*(1), 438–462.

Wilson, S. M., Galantucci, S., Tartaglia, M. C., Rising, K., Patterson, D. K., Henry, M. L., ... Gorno-Tempini, M. L. (2011). Syntactic processing depends on dorsal language tracts. *Neuron*, *72*(2), 397–403.

Woolley, S. (2012). Early experience shapes vocal neural coding and perception in songbirds. *Developmental Psychobiology*, *54*(6), 612–631.

Woolley, S., Hauber, M. E., & Theunissen, F. E. (2010). Developmental experience alters information coding in auditory midbrain and forebrain neurons. *Developmental Neurobiology*, *70*(4), 235–252.

Worsley, K. J., Marrett, S., Neelin, P., Vandal, A. C., Friston, K. J., & Evans, A. C. (1996). A unified statistical approach for determining significant signals in images of cerebral activation. *Human Brain Mapping*, *4*(1), 58–73.

Wu, G., Luo, F., Li, Z., Zhao, X., & Li, S.-J. (2002). Transient relationships among bold, cbv, and cbf changes in rat brain as detected by functional mri. *Magnetic Resonance in Medicine*, *48*(6), 987–993.

Xiong, Y., Zhang, Y., & Yan, J. (2009). The neurobiology of sound-specific auditory plasticity: a core neural circuit. *Neuroscience & Biobehavioral Reviews*, *33*(8), 1178–1184.

Yang, J., & Li, P. (2012). Brain networks of explicit and implicit learning. *PloS One*, *7*(8), e42993.

Yazaki-Sugiyama, Y., & Mooney, R. (2004). Sequential learning from multiple tutors and serial. *Journal of Neurophysiology*, *92*, 2771–2788.

Yu, X., Zou, J., Babb, J. S., Johnson, G., Sanes, D. H., & Turnbull, D. H. (2008). Statistical mapping of sound-evoked activity in the mouse auditory midbrain using mn-enhanced mri. *Neuroimage*, *39*(1), 223–230.

Zhang, N., Rane, P., Huang, W., Liang, Z., Kennedy, D., Frazier, J. A., & King, J. (2010). Mapping resting-state brain networks in conscious animals. *Journal of Neuroscience Methods*, *189*(2), 186–196.

Zhao, F., Wang, P., & Kim, S.-G. (2004). Cortical depth-dependent gradient-echo and spin-echo bold fmri at 9.4 t. *Magnetic Resonance in Medicine*, *51*(3), 518–524.

# List of Figures

# Glossary

**Anterior Forebrain Pathway (AFP)**
A network of nucleï in the songbird forebrain which is responsible for "babbling" behavior in juvenile songbirds and plays an important role in birdsong learning.

**Arcuate Fasciculus (AF)**
White matter tract connecting frontal with parietal and temporal brain regions. Anatomical basis of dorsal language pathway.

**Artificial Grammar Learning (AGL)**
Paradigm used to study acquisition of grammatical rules. In AGL participants acquire grammatical rules from exposure to phrases generated by a grammar and are subsequently tested on their ability to generalize these rules to new phrases.

**Bird's Own Song (BOS)**
Individual song of a zebra finch, in good learners a subtly different copy of the adult song model.

**Blood Oxygen Level Dependent (BOLD)**
Signal on which fMRI measurements are based, representing the fraction of deoxygenized hemoglobin in a brain region. A higher level of deoxy hemoglobin results in BOLD signal loss, the increased BOLD signal associated with neural activation is caused by the increased supply of oxygenized blood to the activated region.

**Broca's area**
Region in the Inferior Frontal Gyrus which has been associated with both speech production and syntactic processing. Broca's area encompasses Brodmann's area's (BA) 44 and 45. In this thesis, Broca's area is seeded by selecting BA 45, because this area has most often been associated with grammar learning.

**Conspecific song (CON)**
      The song of a bird of the same species, either familiar or unfamiliar

**Diffusion Tensor Imaging (DTI)**
      Neuroimaging technique that uses the diffusion of water protons in the brain to visualize white matter tracts.

**Fractional Anisotropy (FA)**
      Scalar value representing the anisotropy of diffusion, where relatively high FA levels are associated with diffusion along an axis, as happens in axonal bundles. Used as a measure of white matter integrity reflecting axonal diameter, fiber density and myelination.

**functional Magnetic Resonance Imaging (fMRI)**
      Neuroimaging technique dependent on the BOLD signal, which is used to measure brain activation.

**Heterospecific song (HET)**
      Song of a bird from another species. In this thesis, song of the European Starling is used as a heterospecific stimulus.

**Language network**
      Set of regions which have been demonstrated to be implicated in language processing and which are functionally connected.

**MLd**
      Dorsal part of the Lateral Mesencephalic nucleus, the primary auditory midbrain nucleus in the songbird brain. MLd is considered the homologue of the human Inferior Colliculus.

**Non-adjacent dependencies (NAD)**
      Dependencies between non-adjacent linguistic elements, often at the level of constituents. In this thesis non-adjacent dependencies exist within three-word phrases which by participants might be interpreted as two constituents, resulting in an adjacent dependency according to linguistic theory. However, the contrast of interest is the presence vs. absence of a dependent rule.

**Resting State Functional Connectivity**
      FMRI technique where the data are collected in the absence of stimulation (in this thesis: with eyes closed) and analyzed for correlations between the low-frequency BOLD signal in different regions. Regions are considered functionally connected when their BOLD signals are correlated under the assumption that "What fires together, wires together".

**Song Control System (SCS)**
      Network of nucleï in the songbird brain which is responsible for the production of song.

**Tutor's song (TUT)**
> Adult song which is memorized by a juvenile bird and used as a model when learning its own song. In the natural environment this is usually the father's song. In this thesis, birds are tutored by other adult males of the same species.

**Vocal learning**
> In this thesis: process in which a juvenile learns its vocalizations by imitating an adult. A criterion for vocal learning is that vocalizations will not develop normally without exposure to adult vocalizations.

# Samenvatting in het Nederlands

Een van de grote uitdagingen voor een pasgeboren baby is het aanleren van zijn of haar moedertaal. Voor een succesvolle verwerving van gesproken taal is het nodig dat baby's (i) aan gesproken taal worden blootgesteld en (ii) dat deze blootstelling vroeg genoeg plaatsvindt, tijdens de zogenaamde sensitieve periode voor taalverwerving. Omdat ook jonge zangvogels aan volwassen zang moeten worden blootgesteld om zelf zang te ontwikkelen en sommige soorten, zoals de zebravink, hun zang alleen binnen een sensitieve periode kunnen leren, zijn zangvogels, en in het bijzonder zebravinken, op dit moment het beste diermodel voor het bestuderen van vocaal leergedrag in andere soorten dan de mens. In dit proefschrift worden zowel zebravinken als mensen met functionele Magnetische Resonantie Beeldvorming (fMRI) bestudeerd om meer te weten te komen over de overeenkomsten en verschillen in de neurale basis van het leren van taal in mensen en het leren van zang in zebravinken.

Aangezien fMRI relatief recent is geïmplementeerd in kleine zangvogels wordt de techniek en de toepasbaarheid hiervan in hoofdstuk 2 beschreven. Vervolgens wordt in hoofdstuk 3 de studie beschreven waarin ik onderzocht hoe zebravinken de zang verwerken die ze tijdens het leren van hun zang van een volwassen mannetje (hun tutor) gekopieerd hebben. Hiervoor krijgen vogels, terwijl er een reeks fMRI beelden wordt opgenomen, zang te horen van hun tutor, zichzelf en een andere zebravink waarmee ze de kooi delen en die van dezelfde tutor geleerd heeft. De resultaten van deze studie laten zien dat de rechter MLd, een auditieve kern in de middenhersenen die homoloog is aan de inferieure colliculus bij de mens, bij volwassen zebravinken nog steeds selectief geactiveerd wordt door de zang van hun tutor in vergelijking met de zang van andere bekende vogels. Bovendien correleerde de amplitude van deze selectieve activiteit met de mate waarin elke vogel de zang van zijn tutor had kunnen kopiëren. Deze bevindingen laten zien dat MLd een rol speelt bij het creëren van een geheugenspoor van de tutorzang tijdens het leren van zang in zebravinken.

Hoofdstuk 4 beschrijft wederom hoe de hersenen van zebravinken tutorzang verwerken, maar in dit hoofdstuk worden de hersenen van jonge vogels bestudeerd om de ontwikkeling van het geheugen voor de tutorzang op de voet te volgen. In dit onderzoek wordt fMRI voor het eerst in jonge zangvogels geïmplementeerd en krijgen de vogels zang te horen van hun tutor, een andere zebravink en een spreeuw. De resultaten laten zien dat ook in jonge vogels MLd selectief geactiveerd wordt door de zang van de tutor. Echter, tegen het einde van de periode waarin jonge zebravinken de tutorzang in hun geheugen opslaan is de linker MLd selectief actief bij het horen van de tutorzang. Bij een vervolgmeting op volwassen leeftijd bleek ook bij deze groep zebravinken de rechter MLd selectief geactiveerd te worden door de tutorzang. Hoewel de groep gemeten jonge vogels klein was, suggereren deze bevindingen dat er tijdens de ontwikkeling van zebravinken lateralisatieveranderingen plaatsvinden in het neuraal substraat van het tutorzanggeheugen. Dit zou een gevolg kunnen zijn van het eindigen van de sensitieve periode voor zangleren.

In tegenstelling tot zebravinkenzang is menselijke taal productief (er is geen limiet aan het aantal mogelijke zinnen en betekenissen), waardoor deze niet geleerd kan worden door imitatie alleen. Woorden en klanken kunnen niet alleen in het geheugen worden opgeslagen, maar ook tot betekenisvolle uitingen worden gecombineerd door middel van grammaticale regels en structuren. In hoofdstukken 5 en 6 wordt een studie beschreven waarin onderzocht werd hoe de hersenen van volwassenen reageren op het leren van de grammaticale regels van een artificiële taal door auditieve blootstelling aan zinnen die aan deze regels voldoen. Volgens de sensitieve periode hypothese zouden volwassenen moeite hebben met het impliciet leren van nieuwe grammaticale regels.

Hoewel de volwassenen inderdaad moeite hadden met het verwerven van de grammaticale regels is zonder vergelijkend onderzoek met kinderen niet te bepalen of dit een gevolg is van het eindigen van de sensitieve periode. De fMRI resultaten lieten zien dat deelnemers wel degelijk geleerd hadden: de selectieve activatie voor de zinnen met vergeleken met zinnen zonder de grammaticale regel die gevonden werd in temporele en frontale taalgebieden, correleerde met het succes waarmee deelnemers later konden aangeven of een zin grammaticaal correct was. Deze correlatie tussen succes bij het leren en hersenactiviteit werd bij de zebravinken in de middenhersenen gevonden. Deze bevinding is een indicatie voor het bestaan van algemene vocale leermechanismen die op verschillende hersen-niveaus tot uitdrukking komen, afhankelijk van het niveau van abstractie dat nodig is om de vocalisaties te leren of te verwerken.

In de hierboven beschreven onderzoeken werd steeds selectieve activatie in specifieke hersengebieden onderzocht. Onze hersenen functioneren echter in netwerken, waarin verschillende hersengebieden met elkaar verbonden zijn. Hoofdstuk 6 gaat in op de relatie tussen functionele en structurele hersen-

verbindingen en het leren van een artificiële grammatica. De resultaten laten zien dat zowel functionele als structurele connectiviteit tussen frontale en temporele hersengebieden die een rol spelen in taalverwerking correleert met de mate waarin deelnemers een artificiële grammatica kunnen leren. Aangezien structurele connectiviteit niet afhankelijk kan zijn van de taak, suggereren deze bevindingen dat de individuele taalontwikkeling een rol speelt in de relatie tussen taalleervermogen en connectiviteit.

Hoewel mensen en zangvogels zowel op hersenniveau en op het niveau van de complexiteit van hun vocalisaties sterk verschillen, doen onze resultaten vermoeden dat er overeenkomsten zijn tussen de neurale mechanismen die vocaal leergedrag mogelijk maken. De studies beschreven in dit proefschrift hebben laten zien dat zowel vocaal leergedrag als de verwerking van geleerde vocalisaties resulteert in selectieve activatie in auditieve gebieden (zebravinken) en gebieden geassocieerd met het abstraheren uit vocalisaties (mensen). Bovendien is de amplitude van deze selectieve activatie gerelateerd aan het individuele leervermogen. Tot slot is bij mensen de functionele en structurele connectiviteit binnen het netwerk dat taal verwerkt ook gerelateerd aan dit individuele vermogen om nieuwe taalstructuren te leren.

# Acknowledgements

A PhD is a journey where the company you bring makes the journey both joyful and succesful. I would like to dedicate this page to everyone who has helped me bring this one to a happy end. First I would like to thank my supervisors, Annemie Van der Linden, Niels Schiller and Claartje Levelt for their support, advice and for their constant belief in my work. I would also like to thank all previous and current members of the Bio Imaging Lab, the LUCL, the LIBC and the Brain and Development Lab for the valuable discussions and suggestions, with a special thanks to the bird team in Antwerp, Eveline Crone and all LACG members in Leiden.

I am grateful to all co-authors of the publications included in this thesis for working with me. A special thanks goes out to Lisbeth and Geert, you are fantastic birdies to work with, and Colline, for always motivating me to do better. Daan, Olga and Ferdi, thank you for being my buddies during those long days at the scanner. I would also like to thank my students, Marloes and Mayke for their dedication and hard work. For helpful advice on task design and analysis I am grateful to Annemarie Kerkhoff and Serge Rombouts. Without support of great technicians and MRI experts this work would not have been possible. Thank you, Johan, Wouter and Mischa. Marieke, thank you for teaching me how to use LaTeX.

Because happy days at the office kept me going I would like to thank my office and coffee mates Martine, Marie-Catherine, Aliza, Yang Yang, Marijn, Bobby and many others for the tea and talks. My friends and family were always there for mental support and distractions when needed. A massive thank you to my parents, my brother and sister and their partners for supporting me and to my close friends (you know who you are) for their friendship and the fun times that helped me to keep calm and carry on. This work would have never been possible without the help of my participants, both human and avian, and I wish to thank them for their participation. I am also very grateful to my paranimphs, Sabrina and Olga, for taking a lot of the stress surrounding the organization of a defense for two universities out of my hands.

# Curriculum Vitae

Anne Marie van der Kant was born May 4th 1984 in Lelystad, The Netherlands. After receiving het atheneum diploma at O.S.G. De Rietlanden in Lelystad, she started her Bachelor degree in Linguistics at Leiden University. Here she first became interested in the sensitive period hypothesis of language acquisition through her thesis research into the syntactic development of congenitally deaf children who received a cochlear implant.

Following her Bachelor degree, Anne moved to Nijmegen for a Cognitive Neuroscience research master, specialization Psycholinguïstics. The research internship for her master thesis was carried out at the Max Planck institute for Psycholinguïstics under guidance of Dr. Anneke Vermeulen and Prof. Dr. Rob Schreuder and resulted in the paper "Reading Comprehension of Flemish Deaf Children in Belgium: Sources of Variability in Reading Comprehension after Cochlear Implantation.", published in Deafness and Education International.

In July 2009, the PhD research described in this thesis was started at the Bio Imaging Lab of the University of Antwerp under supervision of Prof. Dr. Annemie Van der Linden and Dr. Colline Poirier. Results of the zebra finch studies carried out in Antwerp were presented at the 9th International Congress of Neuroethology and the 9th International Conference on the Evolution of Language. Between 2012 and 2014 the human fMRI studies described in this thesis were carried out at the Leiden University Center for Linguistics and the Leiden Institute for Brain and Cognition under supervision of Prof. Dr. Niels O. Schiller and Prof. Dr. Clara C. Levelt. Results were presented at the Society of the Neurobiology of Language 2014 Annual Meeting.

# List of publications

**Full Publications**

**Inter-species comparative research in vocal learning: possibilities and limitations.**
Van der Kant, A.M. In: McCronhon, L, Thompson, B, Verhoef, T, Yamauchi, H (Eds.) *The past, present and future of language evolution research. Student volume following the 9th International Conference on the Evolution of Language*, March 17, 2014

**Functional Magnetic Resonance Imaging (fMRI) with Auditory Stimulation in Songbirds.**
Van Ruijssevelt, L, De Groof, G, Van der Kant, A, Poirier, C, Van Audekerke, J, Verhoye, M, Van der Linden AM. *Journal of Visualized Experiments* 2013;(76), e4369

**Comparisons of different methods to train a young zebra finch (Taeniopygia guttata) to learn a song.**
Dergnaucourt S, Poirier C, Van der Kant AM, Van der Linden AM, Gahr M. *Journal of Physiology-Paris* 2013;107(3): 210218.

**Current state-of-the-art of auditory functional MRI (fMRI) on zebra finches: Technique and scientific achievements.**
Van Ruijssevelt L, Van der Kant AM, De Groof G, Van der Linden AM. *Journal of Physiology-Paris* 2013;107(3): 156169.

**Representation of early sensory experience in the adult auditory midbrain: implications for vocal learning.**
Van der Kant AM, Deregnaucourt S, Gahr M, Van der Linden AM, Poirier C. *PLoS One* 2013;8(4): e61764.

**Reading Comprehension of Flemish Deaf Children in Belgium: Sources of Variability in Reading Comprehension after Cochlear Implantation.**
Van der Kant AM, Vermeulen A, De Raeve L, Schreuder R. *Deafness and Education International* 2010;12(2):77-98

**Conference Proceedings**

**Neural correlates of non-adjacent dependency learning: fMRI and connectivity studies into individual differences.**
Van der Kant A & Schiller N.O. Poster. *Society for the Neurobiology of Language (SNL) Annual meeting*, August 27-29, 2014, Amsterdam, The Netherlands

**Neural correlates of song perception during zebra finch song learning as shown by BOLD fMRI.**
Van der Kant A & Van der Linden A. Poster. *9th International Conference on the Evolution of Language*, March 13-16, 2012, Kyoto, Japan

**Neural correlates of conspecific song perception change during the sensory-motor period in zebra finch males but not females.**
Van der Kant A & Van der Linden A. Poster. *9th meeting of the Belgian Society for Neuroscience*, May 23th, 2011, Leuven, Belgium

**Functional MRI reveals a relationship between the strength of song learning and birds own song selectivity in the auditory midbrain.**
Van der Kant A, Deregnaucourt S, Gahr M, Van der Linden A, Poirier C. Poster. *9th International Congress of Neuroethology*, August 2-7, 2010, Salamanca, Spain

**Reading Comprehension of Flemish Deaf Children.**
Van der Kant AM, Vermeulen AM, Schreuder R, De Raeve L. Poster. *9th European Symposium on Paediatric Cochlear Implantation*, May 14-17, 2009, Warsaw, Poland

**Morphological Development in Dutch-speaking Cochlear Implanted Deaf Children.**
Hammer A, van der Kant A, Coene M, et al. Poster. *13th International Morphology Meeting*, February 3-6, 2008, Vienna, Austria