Neural correlates of vocal learning in songbirds and humans

Kant, A.M. van der

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

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.