University of Groningen

# Topics in auditory fMRI analysis

Renken, Remco Jan

**IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.**

*Document Version*
Publisher's PDF, also known as Version of record

*Publication date:*
2004

Link to publication in University of Groningen/UMCG research database

*Citation for published version (APA):*
Renken, R. J. (2004). *Topics in auditory fMRI analysis: motion-REST-auditory perception and brain response*. s.n.

**Copyright**
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license. More information can be found on the University of Groningen website: https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment.

**Take-down policy**
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

*Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.*

Download date: 11-12-2022

This thesis regards two areas. First, functional Magnetic Resonance Imaging (fMRI) and its analysis, and second, human auditory processing especially with respect to complex tones. Chapter 2 through 4 investigate motion related artefacts in fMRI time series and possible solutions. Chapter 5 and chapter 6 apply the fMRI technique to investigate brain responses to auditory stimuli. Chapter 7 presents psychophysical measurements that investigate the pitch perception of complex sounds.

## Chapter 1

### Sound

In the introductory chapter, chapter 1, the basic concepts of sound and (f)MRI are briefly explained. In this thesis, two types of auditory stimuli are used, pure tones and complex tones. For a pure tone, atmospheric pressure can be described by a single sinusoidal; complex sounds can be constructed by adding several sinusoidal sounds.

There is a relation between the physical parameters and perceptual properties of a sound, although this relation is not always straightforward. The relation between frequency and pitch is relatively simple if a single sinusoidal tone is used; the pitch follows the frequency monotomically. For periodic complex sounds, the pitch percept will mostly correlate to the fundamental frequency.

Sound first reaches the outer ear, then the middle ear, and finally the inner ear where it is transformed into an electrical signal. This transformation takes place at the basilar membrane. Along this membrane a separation in frequency occurs. Nerve fibers, originating at different locations, carry information about different frequencies. However, this separation is not perfect. Two frequencies are often considered to be resolved (distinguishable) if they are more than a third octave apart.

Aside from the frequency information, the auditory nerve also carries temporal information about the temporal envelope of a signal, and about the temporal fine structure.

### (f)MRI

Functional Magnetic Resonance Imaging (fMRI) uses the magnetic properties of protons, to create a three-dimensional image of the brain. In order to measure brain activation, multiple images are recorded and different stimuli/tasks are presented alternatingly. Brain activation is identified by comparing (local) signal intensities between different conditions. The most commonly used method to measure (changes) in brain activity, uses the Blood Oxygen Level Dependent (BOLD) contrast.

There are drawbacks in using fMRI to observe auditory brain response. The rapidly changing magnetic field gradients produce a strong acoustical signal, which will be an extra stimulus for the auditory system. Interference due to this sound can be avoided by using the sparse sampling technique. In this technique there is a long delay between the acquisition of two subsequent volumes. Due to this delay, effects of sound, produced in one volume acquisition, are negligible by the time the next volume is acquired.

### Psychophysics

Auditory psychophysical experiments aim to find the relation between the physical parameters of a sound and their perceptual counterparts. The experiment is mostly done by asking subjects to compare two (or more) stimuli. The perceptual representation is stochastic. If a stimulus parameter is changed gradually, its perceptual representation will change gradually as well. Therefore, a detection probability can often be described with a sigmoidal (logistic) function. The 50% point of this function will indicate the perceptual threshold.

## Chapter 2

In fMRI the aim is to establish (local) changes in image intensity due to the presentation of a stimulus or the execution of a specific task. Any fluctuation of the signal that does not correlate with the task under investigation, is treated as noise. One of the main sources of signal fluctuation is subject motion. This motion can occur within different time windows: 1) during the acquisition of a slice, 2) during the recording of a volume, but in between the recording of two subsequent slices, and 3) between the acquisitions of two successive volumes.[1] In chapter 2 through 4 the focus is on this last type of motion, as this is the most common in fMRI.

During an fMRI measurement, a specific region of an object (or subject) is scanned repeatedly. An inter-volume displacement will affect signal intensities in several ways. First, if an object is displaced, then a different part of the object is scanned. Thus, the same location in an image no longer represents the same location in the object for all scanned images. Second, a displaced object will slightly disturb the main magnetic field. This will cause distortions in the image and alter T2* relaxation, which changes the signal intensity obtained from a certain position in the object. Finally, there is the effect of spin history. When a volume is scanned repeatedly, proton spins within this volume are periodically exited by a radio frequency (RF) pulse. The time between the (start of the) acquisitions of successive volumes (TR) is often not sufficient to allow all spins of the system to return to their equilibrium state; hence a certain amount of saturation will occur. The spin-history artefact is strongest if: 1) motion occurs perpendicular to the slice orientation, 2) the time

---

[1] The term volume refers to a three dimensional image acquired of (part of) the object. The image is constructed by stacking images of slices.

between two volumes is relative short, i.e., $Tr < T_1$, and 3) if there is a gap or there is overlap between two neighboring slices.

In chapters 2 and 3, the effects due to the spin-history artifact are investigated, and a correction procedure for this effect is presented. In order to investigate effects due to spin history, a simulation program was build. This program calculates the magnetic properties of a small volume ($dV$) as a function of time using the Bloch equations. The two main parameters for $dV$ are the longitudinal relaxation time $T_1$ and the equilibrium magnetization $M_{\mathrm{eq}}$.

Using this simulation program we were able to verify that the variation in signal intensity due to a spin-history artefact can be comparable to that generated by the BOLD effect. Therefore it is important to be able to detect and correct for such effects.

An automatic two-step algorithm was designed to detect and correct spin-history artefacts generated by motion. In the first step the occurrence motion and/or spin-history artefacts are identified. In order to do this, the voxels need to be classified as:

- *Null voxels* [N-voxels]: Voxels with an intensity too small to be considered for further analysis;
- *Steady-state voxels* [SS-voxels]: Voxels with a high $T_1$ value relative to the repetition time;
- *Equilibrium voxels* [E-voxels]: Voxels with a small $T_1$ relaxation time compared to the repetition time, causing them to quickly enter into a steady state close to equilibrium between two successive excitations.

This classification is made based on the volumes recorded at the start of an experiment. These volumes contain the necessary information because the system is driven from equilibrium magnetization to a non-equilibrium steady state.

Displaced volumes and voxels affected by spin history can be identified by using the steady-state distribution of voxel intensities of E-voxels and SS-voxels, respectively. After identifying volumes affected by motion and spin-history artefacts, a correction is applied. In the correction step only SS-voxels are corrected. For the affected voxels,the intensity values are replaced by the average intensity of the steady state value, closest in time to the moment of displacement. The correction of the actual displaced volumes is "left" for the regular realignment procedures.

In order to monitor the effectiveness of the correction algorithm, the distribution of the normalized standard deviation (NSD) was used. This value is defined as the ratio of the standard deviation and the temporal average intensity of a voxel. An "ideal" distribution can be obtained by using all volumes that are not affected by spin history. This ideal distribution can then be compared to the NSD distribution of for E-voxels and SS-voxels combined, before and after correction. It should be noted that the voxels/volumens that were displaced were not included in calculating any of these distributions. Using the simulation program, it was shown that the correction step reduced the width of the distribution. Furthermore, the final distribution is similar to the ideal distribution. This shows that the spin-history artefact is effectively reduced.

## Chapter 3

In chapter 3 the motion recognition, and spin-history correction algorithm have been validated using phantom data. To apply precise translations at specific points in time, a computer controlled, movable phantom was designed

The main difference between the simulation program and the phantom measurements is that the simulation program ignores effects due to $B_0$ inhomogeneities, whereas these are naturally present in the phantom measurements. The platform had two possible degrees of motion: translation along the main axis of the bore and rotation around a vertical axis, i.e., perpendicular to the main axis of the bore. Motion always occurred from a pre-set starting position to a pre-set end position. The phantom body allowed for the spatial variation of $T_1$ properties. To test the motion detection and spin-history correction algorithm of chapter 2, two types of motion were applied in the phantom measurements: 1) a single displacement ('out-of-plane'); and 2) a displacement which is immediately followed by the reverse displacement ('back-and-forth'). In general, the results from the phantom measurements support the results using the simulation program. The proposed spin-history correction algorithm substantially decreased the noise induced by the spin-history artefact, and makes its distribution approach the pure noise distribution obtained in the absence of motion.

## Chapter 4

Chapter 4 also deals with the artifacts due to motion, and possible correction mechanisms that have been described in literature. The efficacy of a realignment procedure and subsequent filtering steps have been investigated using: 1) the simulation program presented in chapter 2, 2) phantom measurement with the phantom introduced in chapter 3, and 3) data obtained from two human subjects.

The initial step, in reducing motion artefacts, is the application of a realignment procedure. This procedure consists of two steps: first, estimating the amount of motion (registration) and second, transforming the images (reslicing) using the estimated motion parameters. For fMRI data sets, it is mostly assumed that the head can be considered a rigid body. Therefore, the three rotational and three translational parameters (around /along three orthogonal axes)are estimated for each volume with respect to a reference in the first step. This estimation is often made by minimizing a image-intensity related cost function.

However, even after perfect realignment, there will be residual fluctuations in the fMRI data due to the motion. In order to reduce these residual motion artefacts a filter $Q$ can be applied to the recorded signal($\vec{X}$) for each individual voxel. The filter $Q$ is represented by a matrix. Each row in the matrix contains a filter value for each measured volume. Each column, represents a "subfilter".

In this chapter, four different filters have been investigated:

- $Q_{Frist}$: The filter as proposed by Friston et al. [1996b]. This filter consists of 24 "subfilters": the first six are based on the estimated motion parameters (translation and rotation allong/around the three axis); the second group of six are based on the squared values of the motion parameters. The third and fourth group of six are similar to the first and second, respectively. However they use the estimated motion parameters of the preceding volume.
- $Q_{Bull}$: The filter is constructed as proposed by Bullmore et al. [1999]. Actually a new filter is constructed for each voxel. The motion of a individual voxel can be described by its translation only. Therefore the filter only needs to contain the translation related components of the voxel under investigation. In this case there are 12 "subfilters": the first three based on the translation, the second three based on the squared value of the translation. The third and fourth group are again similar to the first and second, respectively, but they are based on estimated motion of the previous volume.
- $Q_{PCA99}$ and $Q_{PCA95}$: These filters are the result of a principal component analysis on $Q_{Frist}$ retaining those columns that describe at least 99% and 95% of all fluctuations, respectively. This modification has been proposed because we noted a strong correlation between the estimated motion parameters. Furthermore, this will also reduce the number of "subfilters" in $Q$, thereby reducing the number of parameters to be estimated.
- $Q_{True}$: This filter is constructed using the true motion parameters (if these were known), again using the motion parameters and their squared values.

The results from the simulation program and the first subject were used to calculate the spatial distribution of the normalized standard deviation (NSD). This spatial distribution can provide important information when interpreting fMRI data. These distributions clearly show that the highest NSD values are found in regions with a high $T_1$ value, and at the transition between tissues having different $T_1$ and/or proton density. Therefore, structural features of the brain can clearly be observed.

In order to test the efficacy of the realignment step, two different parameters were used. First, the error in estimated motion, and second, the reduction in the NSD distribution. Comparing the measurements/simulations with a long repetition time (TR), i.e., a long time between the acquisition of two subsequent volumes, and short TR, it was striking to see that the error of the motion estimate was generally smaller for the short TR. It appears that if the images are saturated (i.e., there are less details in the image; images become smoother) then the realignment program can estimate the motion parameters more effectively. Another striking phenomenon is that the error in motion estimation followed the overall changes in intensity when a transition from one steady state to another occurs.

Applying any of the filters further reduces the NSD distribution; the extra correction step is capable of filtering motion related signal fluctuations if measurements have been recorded with a short TR. This filtering works especially well in combination with a realignment procedure where the motion parameters are estimated based on the comparisons of image intensities. The errors, made by the realignment program in motion estimation, will have a similar profile as the signal intensity. Therefore, the filter based on these motion parameters is more effective then a filter based on the true motion parameters.

The filters $Q_{Frist}$, $Q_{Bull}$, and $Q_{PCA99}$ are all equally effective in "noise" reduction. Using $Q_{PCA95}$ is also an effective way of reducing motion related fluctuations, but some degradation (compared to the other filters) can be observed. We prefer $Q_{PCA99}$ because it retains more degrees of freedom than $Q_{Frist}$ and it is less computationally intensive than $Q_{Bull}$.

## Chapter 5

In chapter 5, a different part of fMRI data analysis is highlighted. It is shown that the inclusion of a extra control condition, intended to function as baseline, improves data interpretation.

In fMRI, brain functionality is investigated by comparing two brain states: a target and a control condition. Often, the target and control condition are designed such that they only differ in the brain function under investigation. By comparing the target and control conditions, one expects to find a local increase in image intensity during the target condition. However, a positive signal change may not be related to the task. Therefore, a difference found between target and control condition is hard to attribute to either one of the two.

An auditory fMRI experiment is presented in which subjects were asked to compare tones based on their pitches. The stimulus was either a pure tone or one of two complex tones. Each fMRI session (65 scans) consisted of 16 blocks of 4 scans preceded by one scan for synchronisation purposes. The the four scans within a block are labled as S0, S1, S2, and S3 representing the first, second, third, and fourth scan respectively. Each block started with a REST scan, i.e., a scan not preceded by a stimulus. The subsequent 3 scans were preceded by a tone. The time between two scans was 20 seconds, and the time between a stimulus and the scan was 7 seconds. Subjects were requested to report, after the second (S2) and third (S3) stimulus in a block, whether the last presented stimulus had an increased or decreased pitch compared to the previous stimulus (S1 and S2, respectively).

Increased signal intensity was found in several areas, including the left Heschls gyrus and the primary motor and somato-sensory areas, when comparing S2 and S3 to S1. This increase is interpreted as activation and attributed to S2 and S3. Additional areas with a higher intensity during S1 than during S2 or S3 were also found, mainly in the occipital region. One would be tempted to conclude that the increased intensity during S1 is due to a specific task executed in S1. If no other conditions would be present, then this assumption cannot be verified. However, in this experiment a REST condition was included to function as baseline. Using this baseline condition, it is shown that there is a reduction in intensity during S2 and S3, and not an increase during S1. It seems that activity in these areas is suppressed in order to execute the imposed task. In principle, many conditions can function as a baseline, as long as this baseline condition does not require the area of investigation to be (in)active. Knowing whether brain activity increases or decreases during a certain condition is vital to understanding and interpreting fMRI data.

## Chapter 6

This chapter is based on the same data as the previous one, but now emphasis is on the functional description of the auditory fMRI data. The following topics were explored: 1) brain activation as a result of complex- and pure-tone stimuli in relation to their pitch percept, 2) the effect of the direction of pitch change (up or down), and 3) the effects due to the pitch-direction-detection-task executed by the subject. Two complex tones, and one pure tone were used. The complex tones were constructed such that the pitch of the fundamental frequency went up when the frequency of the harmonics went down; the pitch of the pure tone was between the pitches of the fundamental frequencies of the two complex tones.

In this study, several cortical areas showed an increased activation during the processing of complex and pure tones and the execution of the task. Other areas were found to show a decrease in activation. Furthermore, a large variability between the activation patterns of individual subjects was found. The results and implications are summarized as:

- Activation in the left Heschl's gyrus: this is contrary to findings reported in literature. There, activation by tones is either bilateral or dominant on the right side. The activation seen in this study probably has a delayed (maximum) in the BOLD response. Due to the seven seconds delay, between the presentation of a stimulus and the recording of a volume, fast BOLD responses in the auditory cortex (and other areas) have been missed.
- Activation in the lingual gyrus: this result might be an artefact due to its proximity to the straight sinus. If this activation is not an artifact, it may indicate

that different sensory modalities are not fully independent.

- Activation in the inferior frontal region: this is consistent with the memory component of the task. It has been suggested in literature that these regions are related to tonal working memory.
- Activation in the pre- and post-central gyrus: this covers the classical areas for motor and somato-sensory function.
- No areas were revealed that could be indicated as unique for processing pure or complex tones. However, there is some weak indication for areas that are more strongly activated during the processing of pure tones.
- A reduction in the signal intensity in the sensory-motor areas and the occipital lobe: for the sensory-motor area, this decrease occurred on the ipsilateral side to the hand giving the manual response. At the same time, the contralateral motor area showed an increased signal intensity. We hypothesize that, when a response is given with one hand, activity in the ipsilateral motor cortex is reduced to prevent "accidental" movement of the other hands.

  The reduction in the occipital lobe was strongest if a response had to be given. This reduction may be specific for the auditory task, or it may be task independent. In either case we hypothesize that the reduction of activity in occipital lobe regions will reduce —possibly confounding— input to other brain areas that are needed to execute the task at hand. In which case the deactivation can very well be a consequence of regulation of attention.
- In comparing the group results with analysis on an individual level, it was found that the difference between the individuals and between each individual and the group result indicate that there is a large inter-subject variability. The group result, however, do represents the combined, stable results of the individuals; each of the individuals results is not (by definition) a good representative for the group.

## Chapter 7

In chapter 7 the the relation between frequency and pitch is investigated for complex tones using psychoacoustic experiments.

Most periodic complex sounds have a pitch corresponding to their fundamental frequency, $f_0$, which remains audible if no energy is present at this fundamental frequency. A strong pitch percept at the fundamental frequency is heard when the components are resolved; a relatively weak percept can be heard in the unresolved range. Nowadays, the favoured theories on pitch perception assume that both frequency analysis of the resolved components and time-pattern analysis of the unresolved components are involved in pitch perception.

In this chapter the following terms are used:

- *The pitch of the missing fundamental* (short: fundamental pitch), to indicate a pitch related to the fundamental frequency, independent of the underlying mechanism.
- *Spectrally cued pitch* (short: spectral pitch), to signify that the pitch was not at the fundamental frequency. Rather, it correlated to spectral properties of the sound, like spectral edges or mean frequency of the spectrum.
- *Envelope pitch*, which refers to a pitch at the missing fundamental frequency, based on information of temporal envelope modulation, which is present if the stimulus contains unresolved harmonics.

In a 2AFC experiment, subjects had to compare the pitches of two sounds, A and B. Each sound was composed of four successive harmonics of a fundamental frequency between 100 to 250 Hz, added in cosine or in Schröder phase. The harmonic frequencies of A were lower than those of B; the (missing) fundamental frequency of A was higher than that of B. This provides information about whether the subject listened to the fundamental pitch (A higher than B) or to the spectrally cued pitch (A lower than B). The pitch percept was measured as function of $n_A$, the lowest harmonic number of sound A, using a sequential "one up"–"one down" procedure.

For all subjects, the pitch of the missing fundamental dominates the percept at low, resolved harmonics, independent of the relative phase of the components. At the transition from resolved to unresolved components, some subjects showed a transition from fundamental pitch to spectrally cued pitch. For these subjects there was no difference for cosine and schröder phase. However, for most subjects, the range over which the fundamental pitch is perceived extends well into the range of unresolved harmonics when the harmonics are added in cosine phase, while the Schröder phase transition stays within the realm of (partly) resolved harmonics. This shows show that envelope modulation can elicit a relatively strong pitch (stronger than the spectrally cued pitch) even if the stimuli contain only a few harmonics. Overall, the sensitivity to envelope pitch, reflected in the transition point for cosine phase stimuli, was found to vary considerably between subjects.

Our results lead to the conclusion that the auditory system is intrinsically capable of identifying a pitch at the missing fundamental using envelope modulation. However, the extend in which this information is used in the auditory perceptual process is not obvious, as is apparent from the inter subject variability.