Doctoral Thesis

Advanced techniques for functional magnetic resonance imaging of the human brain

Author(s):
Schmidt, Conny F.

Publication Date:
2005

Permanent Link:
https://doi.org/10.3929/ethz-a-005134925

Rights / License:
In Copyright - Non-Commercial Use Permitted
Advanced Techniques for Functional Magnetic Resonance Imaging of the Human Brain

A thesis submitted to the
Swiss Federal Institute of Technology Zurich
for the degree of
Doctor of Natural Sciences

presented by
Conny F. Schmidt
Dipl. Phys. Universität Leipzig
born July 30th 1976
citizen of Germany

accepted on the recommendation of
Prof. Dr. Peter Bösiger, examiner
Prof. Dr. Lutz Jäncke, co-examiner

Zurich 2005
Seite Leer / Blank leaf
## CONTENTS

Summary..................................................................................................................5

Zusammenfassung......................................................................................................9

Motivation, Outline and Significance

Introduction.............................................................................................................13
Sensitivity Encoding in Functional Brain Imaging.................................................14
Spin-Echo Based Functional Contrast.................................................................18
Functional Sequence Design in Auditory Functional MRI.................................19
Conclusion.............................................................................................................21

Specific Theoretical Background

Functional Magnetic Resonance Imaging..............................................................25
SENSE-fMRI...........................................................................................................28
Gradient- and Spin-Echo Based fMRI.................................................................35
Auditory Perception and Acoustic Noise.............................................................39

Chapter 1: Spatial Resolution Enhancement Using Sensitivity-Encoded Echo-Planar Imaging at 3 Tesla in a Typical Motor Paradigm.........................................................47
Chapter 2: Sensitivity-Encoded (SENSE) Echo Planar fMRI at 3T in the Medial Temporal Lobe...........................................................................................................69
Chapter 3: Comparison of fMRI Activation as Measured with Gradient- and Spin-Echo EPI During Visual Perception.................................................................101
Chapter 4: Comparison of "Silent" Clustered and Sparse Temporal fMRI Acquisitions in Tonal and Speech Perception Tasks.........................................................116
Chapter 5: Silent and Continuous Scanning Differentially Modulate Activation in an Auditory Language Comprehension Task..................................................146

Conclusion & Future Prospects.............................................................................169

List of Publications...............................................................................................172
The expanding field of neuroscientific research has brought forward a tremendous amount of knowledge about the brain as the seat of our cognition, emotion, memory, and sensory experiences. Among the multitude of techniques nowadays available to map human brain function, functional magnetic resonance imaging (fMRI) has become the most commonly used non-invasive method over the last decade. Despite its versatility, ease-of-use and widespread availability, fundamental technical obstacles still persist today, which limit the range of its application. Temporal and spatial resolution, signal stability, functional sensitivity, imaging artefacts and acoustic noise are interrelated and can not be improved upon without affecting one another. For example, while temporal resolution can be increased to hundreds of milliseconds, this is only possible compromising on spatial resolution and signal stability, and requires high-performance gradient hardware that gives rise to excessive acoustic noise. Moreover, fast imaging sequences are usually prone to imaging artefacts, such as signal dropout, local distortions and blurring. This thesis aims to surmount problems with susceptibility-related image artefacts and acoustic noise associated with the scanning process. For this purpose, the benefits and constraints of the SENSE technique in comparison to conventional methods are studied, and a 'silent' sequence for fMRI in human audition is implemented and evaluated. The present work was motivated by the specific demands of several collaborative projects in the fields of neuroscience and psychiatry, in which the results were subsequently employed. The scientific context of these studies spans various sensory and cognitive systems in the brain, each requiring specifically tailored acquisition protocols.

Parallel imaging techniques like Sensitivity Encoding (SENSE) use multiple receiver coils with distinct sensitivities for signal detection, which in part contribute to spatial encoding and hence allow accelerating the acquisition process by reducing the number of necessary spatial encoding steps. Thereby they provide a valuable tool to reduce susceptibility-related imaging artefacts, but they can also be tailored to enhance spatial resolution. In particular, the functional specificity and sensitivity associated with SENSE high-resolution functional imaging are investigated in Chapter 1. A typical finger tapping experiment was conducted and the spatial resolution of a conventional single-shot gradient-echo planar sequence was enhanced by a factor of 1.8 in both in-plane dimensions, yielding an effective in-plane resolution of 0.94 x 0.94 mm². Compared to conventional fMRI at moderate spatial
resolution, more spatially detailed and fine structured clusters of activation are clearly apparent, especially in regions with large underlying functional response. However, the potential of high resolution fMRI is limited by statistical sensitivity in regions with low underlying functional activity. These findings might serve for better interpretation of the data when this method is applied in a forthcoming study investigating the functional topography of left and right peri-sylvian areas, that subserve grammatical or prosodic processing.

The application of SENSE provides substantial benefits for reducing susceptibility-related image artefacts; however, intrinsic signal-to-noise-ratio (SNR) deteriorates as a result from undersampling and from spatially varying noise amplification originating from non-unitary operations during the image reconstruction process. The study presented in Chapter 2 therefore aims answering the question as to whether the SENSE technique can be advantageously applied for fMRI at all - especially in the medial temporal lobe, where noise amplification is most problematic. Two fMRI experiments were conducted, comparing SENSE single-shot echo-planar imaging (EPI) at acceleration factors of R = 2.0, 2.4, 2.7, and 3.0 in phase-encode direction with conventional EPI in a face-profession learning task. As expected, susceptibility related image distortions are considerably reduced with SENSE compared to conventional EPI. Yet, acceleration beyond R = 2 reveals only subtle further reduction of geometric distortions, and enhanced encoding speed comes at substantial expense of statistical significance in the functional activation maps, especially in the medial temporal lobe. This result suggests that SENSE acquisition with moderate acceleration (R = 2.0) constitutes a reasonable compromise, that directly entered the acquisition protocols of subsequent studies focusing on the interaction between non-conscious and conscious forms of semantic associative memory, memory consolidation, and on how genes modulate memory function and memory-related brain activity at the behavioural, functional and neuroanatomical level.

Originally initiated by recent and current projects focussing on the functional role of the prefrontal cortex during affective judgement and decision making as well as on the neural representation of distinct emotional dimensions in health and depression (collaboration project 'Cognition and Emotion in Depression: A multimodal functional imaging study'), the third part of the present thesis is dedicated to explore the spin-echo contrast for fMRI in cases where static dephasing is most bothersome. In the study described in Chapter 3, a face perception paradigm was applied in order to compare gradient- and spin-echo based fMRI in
multiple brain regions, and to evaluate which method is best suited to localize activation in prefrontal cortex, where the MR signal is largely dephased in a standard gradient-echo sequence. While activation can be measured with both contrasts, cluster size and significance of the activation as well as the functional signal are significantly smaller in the spin-echo data. Even in the frontal cortex, the spin-echo acquisition did not reveal a noticeable advantage over the optimized gradient-echo sequence (Chapter 2). Consequently the spin-echo approach was not further pursued, and a SENSE gradient-echo EPI protocol is used in consecutive studies that investigate the neural correlates of human altruism and moral judgement.

Using fMRI to study human auditory perception requires the adoption of specifically tailored 'silent' sequences. As part of the project 'short-term and long-term plasticity in the auditory system' a 'silent' clustered sparse sequence was designed and evaluated in two studies. This sequence allows collecting the functional response to the auditory stimulus disentangled from the functional response to the scanner acoustic noise, while, at the same time, it offers time-efficient data collection. In the first study, presented in Chapter 4, the clustered sparse acquisition is compared to a sparse acquisition, showing that the time-efficient data collection is paralleled by enhanced statistical power per time unit. In the second study, presented in Chapter 5, comparison with continuous scanning furthermore revealed that silent fMRI scanning is essential to unambiguously explore the neural correlates of auditory perception and language functions, and that the clustered sparse acquisition is specifically well suited for this purpose. This sequence is now routinely applied in studies that aim to investigate the neural underpinnings of functional lateralization in speech perception as a function of fast and slow prosodic modulations.
Seite Leer / Blank leaf

Parallele Bildgebungstechniken wie SENSE verwenden für die Signaldetektion mehrere Empfangsspulen mit verschiedenen Sensitivitäten, die zum Kodierprozess beitragen und somit durch eine Verringerung der Anzahl notwendiger Kodierschritte den Akquisitionsprozess beschleunigen. Sie stellen ein wertvolles Werkzeug dar, um
suszeptibilitätsbedingte Bildartefakte zu reduzieren, können aber auch zur Verbesserung der Ortsauflösung eingesetzt werden. In Kapitel 1 wird die mit einer hochaufgelösten SENSE-fMRT-Sequenz verbundene Spezifität und Sensitivität untersucht. Ein typisches Fingertapping Experiment wird präsentiert, wobei die Ortsauflösung einer üblichen single-shot Gradientenecho-Sequenz um einen Faktor 1.8 in beiden Raumrichtungen erhöht und eine effektive inplane Auflösung von 0.94 x 0.94 mm² erreicht wird. Verglichen mit üblicher fMRT bei moderater Ortsauflösung werden detailliertere und feiner strukturierte Aktivierungskomponenten sichtbar, speziell in Regionen mit einer starken funktionellen Antwort. Das Potential der hochaufgelösten fMRT ist jedoch durch die statistische Sensitivität in Regionen mit kleiner zugrundeliegender funktioneller Aktivität begrenzt. Die Ergebnisse dieser Studie werden bei der Interpretation der Daten helfen, wenn diese Methode in Folgestudien verwendet wird, um die funktionelle Topographie der linken und rechten perisylvischen Regionen zu untersuchen, die grammatische und prosodische Prozesse unterstützen.

Verhaltens-, funktionellen und neuroanatomischen Ebene erforscht, wie Gene Gedächtnisfunktionen und gedächtnisbezogene Hirnaktivierung modulieren.


Soll die fMRT verwendet werden, um auditorische Wahrnehmungsfunktionen zu untersuchen, bedarf es speziell angepasster 'leiser' Messequenzen. Im Rahmen des Projektes 'Kurzzeit- und Langzeitplastizität im auditorischen System' wurde eine 'leise' clustered sparse sampling Sequenz konzipiert und in zwei Studien evaluiert. Diese Sequenz erlaubt, die funktionelle Antwort auf einen auditorischen Reiz von der funktionellen Antwort auf den Lärm der Messung zu entkoppeln bei gleichzeitig zeiteffizienter Datenaufnahme. In der ersten Studie, welche in Kaptitel 4 beschrieben wird, wird die clustered sparse sampling mit einer herkömmlichen sparse sampling Akquisition verglichen. Die Ergebnisse zeigen, dass die zeiteffiziente Datenaufnahme einhergeht mit erhöhter statistischer Power pro Zeiteinheit. In der zweiten Studie, welche in Kapitel 5 beschrieben ist, zeigt der Vergleich mit einer
kontinuierlichen Akquisition zudem, dass eine leise Datenakquisition essentiell ist, um die neuronalen Grundlagen auditorischer Wahrnehmung und Sprachfunktionen störfrei untersuchen zu können, und dass die clustered sparse sampling Akquisition hierzu gut geeignet ist. Diese Sequenz wird derzeit routinemässig in Studien angewendet, die die neuronalen Grundlagen funktioneller Lateralisierung in der Sprachverarbeitung als Funktion von schnellen und langsamen prosodischen Modulationen untersuchen.
MOTIVATION, OUTLINE AND SIGNIFICANCE

INTRODUCTION

Based on the nuclear magnetic resonance principle, magnetic resonance imaging has evolved into a highly attractive technique for routine clinical investigations and medical diagnostics. Over the past decade this technique was further sensitized to detect local functionally driven signal changes, and hence emerged perhaps the most powerful vehicle in neuroscientific research. Functional magnetic resonance imaging (fMRI) sequences are designed to localize changes in the transverse relaxation properties of functionally activated brain tissue. The functional contrast mechanism results from accompanying changes in blood oxygenation evoked by a specific stimulus as compared to a baseline condition. An ever increasing field of research is now dedicated to develop, integrate and translate new techniques, such as parallel imaging, into practical clinical protocols and sophisticated research applications. Functional magnetic resonance imaging has established a powerful non-invasive and easy-to-use functional neuroimaging approach. In comparison to other functional imaging modalities it readily combines adequate spatial specificity and temporal stability for most applications with a harmless underlying measurement procedure. However, employment of the fMRI technique is circumscribed by the impact of susceptibility-induced image artefacts and signal dropout that result from signal misplacement and dephasing in regions where the underlying magnetic field homogeneity is locally distorted. Moreover, crucial for application in auditory fMRI studies, however, is the acoustic noise associated with the scanning procedure, which might, by acoustic screening and/or psychoacoustic effects, impair the recognition of an auditory stimulus and alter its the perceived sound level and spectral characteristics.

In essence, the present thesis emphasized on two major hot topics in functional magnetic resonance imaging - image quality and acoustic noise reduction during the acquisition process. Recent technical advances in parallel imaging [1] were integrated into routine fMRI practice, and optimized sequence protocols were developed, implemented and evaluated for dedicated fMRI studies in the field of cognitive brain research. Three approaches are presented in order to enhance image quality in different ways, including spatial resolution, image distortion and signal dropout. With regard to acoustic noise a clustered sparse temporal acquisition that disentangles the functional response to the scanner noise from the response to
the acoustic stimulus at issue is evaluated, first with respect to its temporal efficiency in
comparison to a typical sparse temporal acquisition, and second, the impact from acoustic
noise on the pattern of activation. The following paragraphs are intended to put these studies
into their scientific context, to formulate the specific aims and hypotheses, and to place the
results into the framework of previous studies reported in the literature. A detailed description
of the theoretical background follows in a consecutive chapter, and is succeeded by detailed
descriptions of the studies performed in this thesis.

Functional magnetic resonance imaging is based upon the magnetic properties of hemoglobin,
which is paramagnetic in deoxygenated state, but is diamagnetic in oxygenated state.
Deoxyhemoglobin hence behaves like an endogenous paramagnetic contrast agent. The
paramagnetism of the deoxyhemoglobin generates a magnetic susceptibility gradient, which
locally disturbs magnetic field homogeneity that is typically measured by T2*-weighted
sequences. Since oxygenated blood is delivered in excess to the cortical site of activation in
order to replenish the metabolic storehouse of oxygen and glucose, the functional response is
associated with an effective signal increase. Although the specific mechanisms that mediate
between local neural activity and the oxygenation-sensitive T2*-weighted signal are still
controversial, an important empirical link between the electrophysiological signals and the
blood oxygenation level dependent (BOLD) signal has been established [2]. Although
neurovascular coupling was not specifically addressed in the present thesis, it should
nonetheless be pointed out that BOLD-based functional MRI records neural activation only
indirectly with the inherent temporal and spatial resolution of the underlying vascular
response.

SENSITIVITY ENCODING IN FUNCTIONAL BRAIN IMAGING

Previous studies have shown that the functional organization of motor and sensory fields
exhibit a widely distributed and overlapping representation with a gradual somatotopic
gradient [3]. Notably, only very recently, a patchy arrangement of visual, auditory and
multisensory subunits within human superior temporal sulcus has only been unravelled with
high spatial resolution functional mapping [4]. Therefore, distinguished functional patterns of
activation might only emerge at sufficiently high spatial resolution tailored to the underlying
functional topography, and might best be captured by a functional time series that nonetheless
provides appropriate temporal resolution for adequate statistical power. While different approaches have been pursued in the past, such as FLASH sequences or localized partial field of view methods, they all compromise on temporal resolution and whole-brain coverage [5]. In contrast, conventional single-shot echo-planar imaging (EPI) sequences provide sufficient temporal speed to map the whole brain within a temporal resolution of 2-3 s at ample temporal stability, yet at the expense of susceptibility-related image distortions, as will be discussed further down. More importantly, the specific behaviour of the point-spread function needs to be carefully addressed for high resolution functional mapping, since the nominal spatial resolution, as prescribed by the maximum spatial frequency in k-space, might not be achieved in practice as the T2* signal decay essentially acts as a low-pass filter and hence counteracts effective spatial resolution. This in turn results in apparent blurring of the functional images in the spatial domain [6].

In the present work, and explicitly described in Chapter 1, the enhanced encoding speed featured by parallel acquisition, more specifically the Sensitivity Encoding (SENSE) technique [1], is exploited in order to effectively boost spatial resolution in a typical motor experiment. While preserving a temporal resolution of 2 s, the inplane spatial resolution of a standard gradient-echo sequence was increased from 1.6 x 1.6 mm² to 0.9 x 0.9 mm², i.e. by a factor of 1.8 in both in-plane image dimensions. Statistical parametric maps obtained from the high resolution data delineate more spatially detailed and fine structured clusters of activation than conventional fMRI at moderate spatial resolution, especially in regions with large underlying functional response. As a consequence of reduced partial volume effects, clusters of activation in the high resolution data are smaller. Likewise the underlying functional response was larger, as the functional signal is less averaged over non-activated brain tissue. However, the potential of high resolution fMRI is limited by statistical sensitivity especially in regions with low underlying functional signal, since the signal-to-noise-ratio (SNR) is reduced. In order to quantitatively assess the impact from lower statistical sensitivity on one hand and enhanced functional specificity on the other, high resolution data were either spatially filtered in order to reconcile the low resolution behaviour, or the statistical threshold has been adapted to account for different underlying contrast-to-noise-ratio.

In contrast to previous studies that focused primarily on optimal spatial resolution in terms of SNR [7-9], the present study bridges the gap between functional sensitivity and specificity of the functional activation pattern with respect to partial volume effects in an integral manner.
In sum, the data suggest that fMRI can readily be performed with a spatial resolution adapted to cortical functional topography. However, in all potential applications the specific behavior of spatial specificity and statistical sensitivity needs to be taken into account. The present work adds to a roadway to unravel the detailed functional organization of distinguished cortical and subcortical systems. Specifically, current projects focus on the functional segregation within fronto-opercular cortex, as well as on small-scale redistribution and rearrangement of activation patterns underlying experience- and learning-related functional plasticity changes in the motor and auditory system of professional musicians and novice subjects.

In Chapter 2, the SENSE technique is explored as a means to reduce susceptibility-related artefacts. SENSE acceleration not only provides the advantages of enhanced encoding speed in general, and hence reduction of susceptibility-related image distortions in particular, but also a more uniform BOLD-weighting over the acquisition train, since the acquisition is more closely centred around the nominal echo-time. Yet, these desirable potential benefits are accompanied by reduced SNR that results from undersampling k-space on one hand, and from noise enhancement in the reconstruction process on the other. The latter is critical especially in the medial temporal lobe, where noise enhancement is apparently most prominent. Therefore, it is investigated as to whether at all the SENSE technique can advantageously be applied for fMRI in medial temporal lobe structures, such as the hippocampus, which plays an important role for associative memory functions and memory consolidation. Although attempts to reduce susceptibility-related distortions in these anatomical structures have been made earlier in the literature, these focused primarily on image quality, rather than functional activation, and therefore assess image quality in a more descriptive fashion lacking practically relevant alternatives [5]. Techniques relying on gradient compensation, however, can only correct for susceptibility inhomogeneities in specified local regions while affecting other regions oppositely, and importantly, rely on high order shim gradient capabilities [10, 11].

Only two studies so far aimed exploring the encoding speed of the SENSE technique for susceptibility artefact reduction in fMRI [12, 13]. Yet, these are limited to robust functional activation in primary auditory cortex and motor cortex evoked by simple auditory and motor tasks. The underlying functional response were rather strong and deterioration of SNR due to spatially varying noise amplification was not critical. Moreover, in the latter study results are
only presented in a descriptive manner, without critical assessment of the significance of the observed differences related to different experimental protocols.

In contrast, the present study elaborates on experimental protocols in order to provide an optimized setting that is specifically tailored to detect activation in medial temporal lobe structures. This particular region is in typical echo-planar acquisitions severely affected by susceptibility-related image distortions, and hence serves a good candidate to benefit from SENSE acceleration, but on the other hand is critical in terms of noise amplification from the SENSE reconstruction. Two fMRI studies are carried out in order to assess the effect from different imaging parameters. In the first part a conventional EPI sequence is compared to a SENSE-EPI acquisition with a SENSE acceleration of R = 2 at two echo times, TE = 22 ms and TE = 35 ms. The echo-time of the EPI sequence sets the total time scale for intravoxel dephasing. Since this goes along with a concurrent reduction of T2*-weighting, a compromise with respect to optimal BOLD sensitivity is to be made. In the second part benefits and drawbacks from higher SENSE acceleration are explored using SENSE accelerations of R = 2.0, R = 2.4, R = 2.7 and R = 3.0.

A paired-associate learning paradigm was disposed in order to evoke activation bilaterally in the medial temporal lobe and furthermore in the inferior frontal and the left and right fusiform gyrus. Learning success was assessed with a cued recall task immediately after the functional scan. The retrieval performance was included as a covariate in the statistical comparison between acquisition protocols in order to partial out variance underlying inter- and intra-individual cognitive differences. As expected, SNR is significantly reduced between the data acquired at TE = 22 ms and TE = 35 ms and signal is almost completely dephased in typical susceptibility-affected regions, such as orbitofrontal cortex. With increasing SENSE acceleration, SNR progressively drops and noise amplifies in the medial temporal lobe. It is of interest, however, that SNR reduction from SENSE acceleration is less severe than predicted from theory. This eminent finding is interpreted in the framework of recent models of physiological noise contributions [12, 14]. Most importantly, the functional images show that susceptibility-related image distortions are markedly reduced with the SENSE as compared to the conventional EPI acquisition. Data acquired with SENSE acceleration factors beyond R = 2 reveal only subtle further reduction of geometric distortions, and enhanced encoding speed is at remarkable expense of statistical significance in the functional activation maps, particularly in the medial temporal lobe. Yet, an unexpected improvement of functional sensitivity emerged at extreme acceleration factor (R = 3.0), which, eventually, might
originate from residual signal foldover, an effect that has not attracted interest in previous studies so far. Taken together, the data from this study suggest that with a SENSE acquisition susceptibility-related image distortions can be markedly reduced and a moderate SENSE acceleration factor of $R = 2.0$ is appropriate for functional studies aiming at the detection of medial temporal activation. Experimental protocols were subsequently adopted, permitting studies that probe on the interaction between non-conscious and conscious forms of semantic associative memory, examine the neural correlates of memory consolidation, and explore how memory function and memory-related brain activity is genetically predisposed [15, 16].

**SPIN-ECHO BASED FUNCTIONAL CONTRAST**

The BOLD signal measured with $T_2^*$-weighted gradient-echo EPI is composed of a contribution from a static dephasing and dynamic effects. The static component emerges from magnetic field perturbations which increase the heterogeneity of the phase distribution across a voxel. In contrast, the dynamic component results from irreversible loss of phase coherence due to diffusion of water molecules in the susceptibility gradient of paramagnetic deoxyhemoglobin. The functional contrast obtained with spin-echo based fMRI hence differs from the contrast that underlies gradient-echo based fMRI in so far, as in a spin-echo experiment static dephasing effects that contribute to the BOLD signal are refocused. This results in reduced functional signal by a factor of 2 - 4, depending on the underlying vascular architecture. Although providing lower functional signal, spin-echo based fMRI can be beneficial in regions where signal is largely dephased in a gradient-echo acquisition. This led to the hypothesis that in regions that suffer from susceptibility-related dropout, such as orbitofrontal cortex, the MRI signal can be restored with a spin-echo sequence and hence a functional signal detected.

Part of the present thesis is dedicated to explore the spin-echo contrast for fMRI in cases where static dephasing is most bothersome (Chapter 3). Previous studies have shown that spin-echo based fMRI is a useful tool to map functional activation in primary visual and motor cortices [17, 18]. Moreover, it has been shown that spin-echo contrast at 3 Tesla can successfully be applied to cognitive fMRI studies [19]. In the present study activation measured with gradient- and spin-echo EPI sequences is directly compared during visual perception. A simple task, namely attentive viewing of faces in various stimulus formats (line
drawings of unfamiliar faces and photographs of unfamiliar, famous, and emotional faces), was used in order to localize activation within multiple, bilateral face-responsive regions in extrastriate cortex, the limbic system, and prefrontal cortex. With both gradient- and spin-echo EPI sequences activation is localized in the same face-responsive regions, including the inferior occipital gyrus, fusiform gyrus, superior temporal sulcus, amygdala, inferior frontal gyrus, and orbitofrontal cortex. However, the spatial extent of the activation, statistical significance, and amplitude of the fMRI signal are significantly smaller in the spin-echo EPI data. Even in the orbitofrontal cortex the spin-echo acquisition has not revealed a significant advantage over the gradient-echo acquisition in terms of functional signal and significance. Although signal was restored in this typical dropout region, the inherently lower functional contrast of spin-echo based fMRI did not reveal stronger activation when compared with gradient-echo based fMRI. The spin-echo approach only performs better than gradient-echo based fMRI in regions where its inherently lower functional contrast outweighs the signal (and hence functional signal) loss associated with the static dephasing in the gradient-echo sequence. These data therefore indicate that optimized gradient-echo sequences that reduce susceptibility artefacts (as elaborated in Chapter 2) are sufficient to detect activation in the orbitofrontal cortex. The spin-echo approach is hence not further pursued, and a SENSE gradient-echo EPI protocol is used in several studies performed in the context of the collaboration project 'Cognition and Emotion in Depression: A multimodal functional imaging study' [20]. It is furthermore continuously applied in novel projects, which aim investigating the neural correlates of human altruism and moral judgement.

FUNCTIONAL SEQUENCE DESIGN IN AUDITORY FUNCTIONAL MRI

The second focus of the present thesis is directed to the development and implementation of optimized sequences for functional magnetic resonance imaging in the auditory system. Customary functional sequence protocols are hampered by acoustic noise generated during operation of the MR system, which affects the experimental integrity of auditory studies [21, 22]. Over the past decade important advances have been made in terms of acoustic noise reduction, including passive and active shielding, hardware engineering design that reduces generation and transmission of acoustic noise e.g. by Lorentz-force balanced gradient coils, acoustically damped mounting system and gradient support structures, and selection of gradient waveforms that avoid acoustic resonances and high-frequency noise (for review see
However, these prospects might only reduce, but not eliminate acoustic noise completely. Therefore, functional designs have been adopted that optimally exploit the sluggishness of the functional response. The functional response to the stimulus at issue can then be disentangled from the functional response to the scanner acoustic noise [25]. Yet, temporal efficiency of such designs is of critical importance, since data are only sparsely acquired, which culminates in lengthy experiments for ample statistical analysis. The two studies described in detail in Chapters 4 and 5 present and evaluate a newly devised and implemented functional clustered sparse temporal design. Chapter 4 focuses on the temporal efficiency of the new sequence compared to a sparse temporal design that has previously been described in the literature [26]. The clustered sparse temporal acquisition is then applied in Chapter 5 as a 'silent' acquisition scheme in a subsequent study that assesses the impact of ambient scanner acoustic noise on the functional activation pattern by way of comparison with a continuous scanning approach.

Specifically, the first study directly compares and examines the functional response measured with the clustered and sparse temporal acquisition schemes by way of both a tonal and language comprehension task. A significant functional response is localized in regions known to mediate auditory perception and language processing and clear advantage is demonstrated for the clustered as compared to the sparse temporal acquisition. With the clustered acquisition larger and more significant clusters of activation are obtained consistently in all subjects and within all regions of interest. These data show that the clustered temporal acquisition allows for time-efficient data collection which is paralleled by enhanced statistical power per time unit. This result implies, that the gain in statistical power might conversely be sufficient to reduce the number of trials required to perform a reliable statistical analysis, and therefore a reduction of total scan time. In sum, this study demonstrates that the clustered sparse temporal acquisition provides a valuable and effective tool, which is particularly tailored to the specific demands of auditory fMRI, and at the same time provides enhanced sampling efficiency. Hence it might help increase current knowledge on the functional architecture of auditory perception and to identify neural pathways subserving specific auditory and language functions in a silent scanning environment. These amendments might furthermore help improving the applicability of event-related functional MRI in clinical settings, as in particular patients may benefit from this comfortable scanning procedure which nonetheless provides sufficient statistical power to obtain powerful statistical maps from individuals. Finally, it has been shown, for the first time, that this novel approach is evidently
applicable in studies examining the level of spoken stimuli longer than one word or two, which opens a new horizon in the realm of neurocognitive research of speech and language.

In order to assess the impact of scanner acoustic noise and from the number of data acquired during each event-related functional trial, the pattern of activation evoked by a language comprehension task is explored by way of a 'silent' and continuous acquisition. The specified results demonstrate that activation in primary and secondary auditory cortices is modulated by neuronal habituation and/or hemodynamic saturation in the presence of continuous acoustic sensory input, and thereby exhibits a larger response in the 'silent' acquisition during the same task. In contrast, regions that mediate higher language functions are more engaged in the continuous acquisition, since enhanced effort is necessary to retrieve intelligible information from the stimuli in the noisy environment. Furthermore, it is demonstrated that ambient scanner acoustic noise not only modulates functional activation oppositely in distinct regions of the brain, but has also an effect on the lateralization of the functional response. This result is supported by updated models of the functional organization of the auditory system that ascribe different acoustic cues to relatively specialized lateralized processing domains [27]. It is argued that temporal acoustic cues preferentially recruit the left hemisphere, while spectral cues are processed more efficiently in the right hemisphere. Since auditory stimuli in a noisy environment are perceived as relatively flat sounds, the right hemispheric response might hence relatively be damped during continuous scanning. In conclusion, these findings implicate, that silent fMRI scanning is essential to unambiguously explore the neural correlates of auditory perception and spoken language. And the clustered sparse temporal acquisition is specifically tailored for this purpose, since it combines silent fMRI scanning devoid of interference with scanner acoustic noise and time-efficient data collection. It is now routinely applied in current projects elaborating on the neural underpinnings of functional lateralization in speech perception as a function of fast and slow prosodic modulations.

CONCLUSION

The proposed methods and preferential protocols that were developed and evaluated in the present work provide an improved access to study the topographical organization of human brain function as well as the mental processes that mediate human adaptive behaviour and social interaction. Specifically, the selected measurement protocols are conceived for specific
application in complex neuroscientific investigations. The implications from each separate part of this work were readily integrated into distinguished collaboration projects, where they are effectively used in promising application studies. The emergent emendations from the present thesis provide a powerful approach to address specific problems associated with susceptibility-induced artefacts and distortions in the functionally sensitized echo-planar images, and the acoustic noise generated during the acquisition process. They might furthermore contribute to the enhanced realization of future studies.

REFERENCES


FUNCTIONAL MAGNETIC RESONANCE IMAGING

Over the last decade functional magnetic resonance imaging (fMRI) has emerged as the most often used technique for mapping human brain function. Almost all sensory and cognitive systems have been studied with this technique in health and disease, in order to explore the functional topography of our brains and changes in functional organization during development and learning, but also to learn about the neural correlates of functional abnormalities and functional reorganization after brain damage. FMRI allows to non-invasively detect subtle changes in blood oxygenation that are driven by brain activation, and the functional contrast is hence termed blood oxygenation level dependent (BOLD) contrast [1]. The BOLD signal is measured by the effects of local magnetic susceptibility field gradients that are induced by local alterations in deoxyhemoglobin content.

Figure 1: Functional mapping methods

picture taken from: http://www.fmrib.ox.ac.uk/~peterj/lectures/miccai99/sld002.htm

Figure 1 illustrates how fMRI ranks among other functional mapping methods in terms of spatial and temporal resolution and invasiveness of the technique. Single cell recordings allow to record synaptic transmission of action potentials at their very source with high spatial and temporal precision. In contrast, fMRI provides a tool to study networks of neuronal populations that interact with each other over large scale distances across the whole brain.
Although fMRI can not compete with the temporal resolution provided by EEG (electroencephalography) and MEG (magneto-encephalography) or single cell recordings, it is not invasive like the latter and spatially more precise than the former. In contrast, PET (positron emission tomography) permits to selectively study distinct brain metabolites, however at a longer timescale and with less spatial specificity. Hence, fMRI turned a method of choice to study human brain function, enabling comparatively high spatial and temporal specificity while being a non-invasive, widespread available, versatile and harmless procedure.

It is important to note that the functional signal measured with fMRI only indirectly measures brain activation and originates from the combined effect of distinct vascular events and metabolic changes (Figure 2). Neuronal activation is accompanied by a release of one or more mediators in the central neuronal layers, for example NO, K, and/or adenosine, which induce a smoothing of the muscle cells on the arterial side and subsequently cause the blood vessels to dilate, which in turn triggers an increase in blood flow and volume. However, the precise mechanism that mediates between electrical activity and the hemodynamic response as well as the sequence and interplay between metabolic, physiological and hemodynamic events are not yet fully understood. Various models have been brought up to describe activity-induced coupling between cerebral blood flow and oxygen consumption, spanning uncoupling [2, 3], transient uncoupling [4, 5] and tight coupling [6, 7]. Even at a biochemical level the exact mechanism that drives synaptic glucose metabolism is rather controversial. The Astrocyte-Neuron Lactate Shuttle Hypothesis proposed by Pellerin & Magistretti [8, 9], which states that astrocytes provide lactate as an energy substrate for neurons, has recently been critically reviewed by Chih & Roberts [10], supporting the conventional hypothesis, which contends that activation-induced energy demand is met predominantly by metabolizing glucose oxidatively [11, 12]. Moreover, the very coupling mechanisms as well as the vascular reactivity supposedly differ between brain regions and also in health and disease. Fortunately, it has recently been proven that the BOLD signal indeed tightly correlates with change in local field potentials in postsynaptic neurons that occurs as a result of changes in excitatory (glutamatergic) neurotransmission [13].
An fMRI experiment essentially includes the acquisition of a time series of susceptibility (T2*)-weighted images during which the subject is presented with a specific task that alternates with a baseline state. Further, a statistical map is generated by analysing the functional time series with respect to correlation between the induced signal changes and the course of the applied functional task. Finally, an anatomical scan is acquired for structural reference, as outlined in Figure 3. For the interpretation of the resulting data it should precisely be indicated what is actually meant by 'activation', since it might be quantified by either the size or significance of the observed clusters of activation, or the amplitude of the BOLD signal change.
SENSE-FMRI

In order to describe the very foundations of parallel imaging and to elucidate its potential benefits and drawbacks, the following section is dedicated to the mathematical underpinnings of the image reconstruction process.

The principle of nuclear magnetic resonance (NMR) as used in MRI is that hydrogen nuclei are electrically charged and have an inherent angular momentum or spin $\mathbf{l}$, that generates a magnetic momentum $\mu = \gamma l$, $\gamma$ being the gyromagnetic ratio. This angular momentum is quantized, and according to their spin quantum number $I = 1/2$, hydrogen nuclei arrange in two energy states or eigenstates in an external magnetic field. Transitions between these two states are induced by an electromagnetic field with a frequency $\omega$, when the energy transfer to the system matches the transition energy $\Delta E = h\omega (\hbar$ Planck's constant) between these two states. When the spins return to their original state, they emit an electromagnetic wave with the same frequency, usually called the resonance frequency, forming the MR signal.

Spatial encoding of the MR signal is based on the fact that the resonance frequency is dependent on the magnetic field strength $B_0$, i.e. $\omega_0 = \gamma B_0$. Hence, if additional spatially dependent magnetic fields (so-called imaging 'gradients' $G = \nabla B$) are applied to the object, the MR signal can be localized by its frequency. Based on this principle different techniques have been developed. The first among these were based on slice selective excitation [14] and the projection-reconstruction method, whose image reconstruction procedure closely resembles the image reconstruction algorithms of computer tomography [15].

To date the most widely used method is the Fourier imaging method [16]. After a slice selective excitation and prior to the actual acquisition of the MR signal, a so-called phase encoding gradient $G_y$ is applied during a time interval $T_y$. According to their position along the $G_y$ gradient field, the spins will precess with different frequencies ($\omega_y = \gamma (B_0 + yG_y)$) and thus show different phases at the end of the time interval $T_y$. Hence the y-coordinates of an object are 'phase-encoded'. The MR signal is then acquired during a time interval $T_x$ under a gradient $G_x$, and hence the x-coordinate is encoded in the Larmor frequency ($\omega_x = \gamma (B_0 + xG_x)$). In other words, a two-dimensional image is spatially encoded along a phase and

---

1 FID = free induction decay of the signal after the excitation pulse
frequency domain, which span the Fourier space or so-called k-space, and the k-space matrix is filled by sampling at consecutive time points in x-direction, while the time is fixed in y-direction and instead the gradient is varied (Figure 4).

The image of the object can then be reconstructed from the acquired k-space data by Fourier transformation, which reads:

\[ m(x, y) = \int \int M_T(k_x, k_y) \cdot \exp(i(k_x x + k_y y))dk_x dk_y \]  \[1\]

- \(m(x, y)\) signal of the object in image space
- \(M_T(k_x, k_y)\) signal of the object in k-space
- \(k_{x,y} = \gamma G_{sij} t\) k-space coordinates

It is this very relationship between the signal in k-space and in image space that we need to keep in mind, when we want to understand certain physical features of the reconstructed image and image artefacts (Figure 5). It should further be noted, that it follows directly from this mathematical relationship that every data point acquired in the k-space contributes to every pixel of the final reconstructed image.
**Figure 5:** Relationship between the Fourier space ('k-space', in which the data are acquired in) and the image space. 

a. Images are reconstructed by Fourier transformation of the acquired k-space data. 
b. The spacing of sampling points in the k-space codes for the size of the image, i.e. the image field of view (FOV). Therefore, if only every \( n \)th line of the k-space data matrix is used for reconstruction ('undersampling'), the image field of view (FOV) is \( n \) times smaller, and moreover the reconstructed image will be \( n \) times folded. 
c. The size of the sampling matrix, on the other hand, codes for the resolution of the reconstructed image (\( \Delta x \)). If only the central points of the k-space data matrix are sampled, the reconstructed image only contains low-level features of the imaged object at low spatial resolution. 
d. In contrast, data obtained from the periphery of the k-space data matrix only represent the fine-structured features of the object.

**Parallel Imaging and SENSE**

The basic idea of parallel imaging is to use arrays of multiple receiver coils with distinct spatial sensitivities. The MR signal is sampled parallel (simultaneously) by the different coils, that also contain distinct information about the spatial distribution of signal sources in the sample. Hence gradient encoding can partly be made redundant and encoding steps may be skipped, which in turn enhances encoding speed. Since sparing of phase-encoding steps in k-space results in a decrease of the field of view (Figure 5b), the image of the object is folded ('aliased'). Yet, the full field of view of the image can be reconstructed by using spatially varying coil sensitivity profiles (Figure 6). The maximum acceleration factor \( R \), by which
sampling in k-space can be reduced, is basically determined by the number of independent coils, however in practical applications is limited by the SNR penalty resulting from undersampling and noise enhancement in the parallel imaging reconstruction. A recent study showed that at higher field strengths the transition between favourable and prohibitive parallel imaging conditions shifts toward higher feasible acceleration factors, which makes high field MRI especially attractive for parallel imaging applications.

While SMASH (Simultaneous Acquisition of Spatial Harmonics, [17]) algebraically combines the measured k-space data such as to synthesize the missing information in k-space, SENSE (Sensitivity Encoding, [18]) operates in image space, and uses the spatially varying coil sensitivity profiles to disentangle signal components that originate from distinct parts of the object and to reconstruct the full field of view of the image.

**Figure 6**: Parallel imaging and Sensitivity Encoding (SENSE). In contrast to a full Fourier-encoded acquisition, parallel imaging allows to spare phase encoding steps ('undersampling') and hence to speed up the encoding process. The resultant aliased image is unwarped by using the spatially varying coil sensitivity profiles to disentangle signal components that originate from distinct parts of the object.
SENSE and High Spatial Resolution

In principle, high spatial resolution can be achieved by simply increasing the size of the sampled k-space matrix, resulting in an increase of the acquisition time. This, however, makes it impractical in the case of echo-planar imaging (EPI), where the whole k-space matrix is acquired following one excitation pulse ('one shot'). During the course of the acquisition the MR signal fades with a time constant that is determined by the properties of the underlying tissue (i.e., the transverse relaxation times). For example, in a gradient-echo sequence signal from gray matter decreases with a half-life time of $T_{2^*} = 42$ ms at 3 Tesla field strength. This very signal decay essentially acts as a low-pass filter in k-space, because the data acquired at the periphery of the k-space do no longer encode significant signal from the object but rather underlying noise and hence do not anymore resolve fine-structured detail of the object (situation similar as in Figure 5c). That means that although the nominal spatial resolution as imposed by the size of the k-space matrix is still high, the object will appear blurred, as illustrated in Figure 7. In mathematical terms this situation can be described by the full widths at half maximum (FWHMs) of the sampling point-spread function (PSF) and the broadening of the PSF by the $T_{2^*}$-decay, which is described by a Lorentz function.

Therefore, in order to make advantage of high spatial resolution, one needs to operate in a regime where the line width of the Lorentz function is smaller than or equal to the nominal resolution imposed by the sampling PSF. This can be achieved by reducing the overall acquisition time, and hence the echo train length. If single-shot EPI is not imperative, k-space data might be sampled in segments (multi-shot EPI), which prolongs total imaging time for one volume. However this might not be affordable when functional time series are to be obtained with high temporal precision and when whole brain coverage is mandatory. Additionally, inconsistencies between consecutive 'shots' might lead to image artefacts. It hence becomes clearly apparent that parallel imaging provides a far more elegant solution in this situation. Rather than sampling of the full k-space matrix, total acquisition time is reduced by sparing acquisition steps in the phase-encoding dimension (Figure 7). However, it should finally be mentioned that the intrinsic spatial resolution of the BOLD contrast itself sets the ultimate spatial resolution limit for fMRI, which is essentially determined by the anatomy of the venous vasculature.
Figure 7: The nominal spatial resolution of an image can be enhanced by increasing the size of the k-space matrix. However, the effective spatial resolution might be lower when signal decay over the course of the data acquisition causes a line broadening of the PSF that exceeds the nominal resolution imposed by the size of the k-space sampling matrix, resulting in a blurred image. However, with the SENSE technique the echo-train length and hence line broadening due to signal decay can be reduced and high spatial resolution indeed be achieved.

**SENSE and Susceptibility Artefact Reduction**

Parallel imaging in combination with single-shot EPI might not only be tailored to provide enhanced spatial resolution in the same total acquisition time as a conventional full-Fourier acquisition, but can on the other hand be used to shorten the acquisition duration at preserved spatial resolution.

Since susceptibility gradient fields across the brain add to the imaging gradients, the measured MR-signal in k-space is not necessarily associated to the coordinates defined by the imaging gradients only, but it is mislabeled as is described by a susceptibility term

\[ \exp(-i\gamma \delta B t) = \exp\left[-i \left( k_x \frac{\delta B(x, y)}{G_x} + k_y \frac{\delta B(x, y)}{G_y} \right) \right] \]

that needs to be included in Eq. [1] for the
reconstruction of the true image. Furthermore a term $\exp(-t/T_2^*)$ needs to be added, that describes the $T_2^*$ decay of the signal. Eq. [1] then reads:

$$m(x,y) = \int \int M_T(k_x,k_y) \cdot \exp(i(k_xx+k_yy)) \cdot \exp(-i\gamma B_t) \cdot \exp(-t/T_2^*) dk_x dk_y$$  \[2\]

$\exp(-i\gamma B_t)$ susceptibility field $B$ that adds to the encoding process

$\exp(-t/T_2^*)$ $T_2^*$ decay

Hence a particular image feature might be misplaced after standard Fourier reconstruction of the k-space data. In the worst case, the MR signal is even not acquired at all when susceptibility gradients shifted it outside the prescribed k-space matrix. Yet it is clear that if the acquisition time $t$ is small, then the additional susceptibility term is close to 1, and the standard Fourier reconstructed image approaches the true image of the object. In other words, the shorter the acquisition time, the less susceptibility-related image artifacts can evolve.

Under the simplifying assumption of linear susceptibility gradient fields (for simplicity only the phase-encoding direction $y$ is considered):

$$\delta B = \alpha + \beta_y y$$

the signal in k-space originating from the signal $m(y)$ at position $y$ in the object translates into:

$$M_T(k_y) = \int m(y) \cdot \exp(-i\gamma G_y (y + \alpha / G_y + \beta_y / G_y y) t) dy$$

This equation can be transformed into a standard Fourier transformation by variable substitution

$$y' = y(1 + \beta_y / G_y) + \alpha / G_y$$

\[2\] Same equations apply for the readout direction $x$, however the effects are less severe than in phase-encoding direction, since the acquisition bandwidth is larger in readout direction.
\[ M_r(k_y) = \frac{1}{1 + \beta_y / G_y} \int m(y') \cdot \exp(-iG_y y') dy' \]  

Hence, a susceptibility field \( \delta B \) might cause the intensity of the image to change by a factor of \((1 + \beta_y / G_y)^{-1}\), but also to shift the image by \( y' - y = y\beta_y / G_y + \alpha / G_y \), and to magnify the image by \( dy'/dy = 1 + \beta_y / G_y \).

The application of SENSE therefore provides substantial benefits for reducing susceptibility-related image artefacts, simply by reducing the required acquisition time - and in turn application of larger imaging gradients. However, intrinsic signal-to-noise-ratio (SNR) deteriorates, which on one hand results from undersampling itself (since less data are acquired) but on the other hand is evoked by spatially varying noise amplification, which results from non-unitary operations during image reconstruction (and is quantified by the so-called g factor [18]). Therefore the question raises as to whether at all the SENSE technique can be advantageously applied for fMRI. Part of this thesis aimed answering this question with a special focus on the medial temporal lobe, were the g factor is most problematic.

**GRADIENT- VERSUS SPIN-ECHO BASED FMRI**

*Static and dynamic dephasing effects*

The BOLD signal arises from the paramagnetism of venous blood deoxyhemoglobin, and functionally driven alterations in blood oxygenation (and hence a change in deoxyhemoglobin content) are measured as a BOLD signal change. The paramagnetic deoxyhemoglobin gives rise to a local magnetic susceptibility field gradient intravascularly, but also between the venous vessel and the surrounding diamagnetic tissue.

Two different spin-dephasing mechanisms contribute to the BOLD signal (Figure 8), dynamic (diffusion) effects \((\sim B_0^{-2})\) and static dephasing effects \((\sim B_0)\). The static dephasing effect arises from loss of coherence of tissue water spins in the susceptibility gradient field between the intra- and extra-vascular compartment. The dynamic effect is based upon diffusion of tissue water spins in the susceptibility gradient field around blood vessels, and diffusion of...
blood water spins in the susceptibility gradients around deoxygenated red blood cells within the blood vessels (intra-vascular BOLD effect as described further down).

The so-called transverse relaxation times $T_2$ and $T_2^*$ are the half life times that describe the decay of the signal due to these spin-dephasing effects. $T_2^*$ is defined by the following equation:

$$\frac{1}{T_2^*} = \frac{1}{T_2} + \frac{1}{\tau}$$

where $\tau$ represents the relaxation effect evoked by static field inhomogeneities. While dynamic effects are measured with spin-echo (T2-weighted) weighted sequences, both dynamic and static dephasing effects contribute to signal changes measured with gradient-echo ($T_2^*$-weighted) sequences.

Figure 8: Two different spin-dephasing mechanisms that contribute to the BOLD signal. GE = gradient-echo, SE = spin-echo experiment.
**Extra-vascular and intra-vascular BOLD-effect**

The extra-vascular BOLD effect arises from both static dephasing and diffusion effect experienced by tissue water spins in the susceptibility gradient field around the blood vessels. If the range of the susceptibility gradient field around the vessel is larger than the diffusion length, static dephasing effects contribute more to the signal loss than diffusion effects (Figure 9). Typically, small vessels show mostly diffusion losses.

The intra-vascular BOLD effect arises from the diffusion of blood water spins in the susceptibility gradient field around deoxygenated red blood cells within the blood vessels and is characterized by rapid exchange of water molecules between red blood cells and blood plasma.

Not only does the susceptibility difference between the intra- and extra-vascular space change with blood oxygenation – and hence activation, but also the transversal relaxation time $T_2$ of blood itself is dependent on oxygenation (Lutz-Meiboom Modell, [19]). The latter can give rise to a BOLD signal in, and the former around large veins, remote from the actual site of activation, when changes in blood oxygenation propagate downstream in the venous vascular bed. Hence the intra- as well as the extra-vascular BOLD signal arising from larger veins needs to be eliminated in order to map functional activation with high spatial specificity.

**Spin-echo and gradient-echo based fMRI**

The functional contrast obtained with spin-echo (SE) based fMRI differs from the one underlying gradient-echo (GE) based fMRI. In an SE experiment, static dephasing effects that contribute to the BOLD signal are refocused, which results in reduced functional signal by a factor of 2–4 [1], depending on the underlying vascular architecture. Furthermore, the same static dephasing effects give rise to signal dropout in regions like the orbitofrontal cortex, where the MR signal is dephased in the magnetic field inhomogeneity generated between the air-filled cavity of the frontal sinus and the surrounding brain tissue. It was investigated in this thesis, whether in this very region the MR signal could be restored with an SE acquisition and a functional signal be detected. Although effects of compartmentalization and spatial specificity and selectivity of GE- and SE-based functional contrast mechanism (Figure 9) were not specifically addressed, they should be mentioned in the following, since the 'brain-
vein-problem constitutes one of the most pertinent methodological fMRI issues. Several properties of the BOLD signal, however, might help to surmount this problem.

GE-based fMRI

\[
\frac{1}{T_2} = \alpha \cdot (\Delta x \cdot (1 - Y) \cdot \varepsilon) + b_0 \quad \text{large vessels}
\]

\[
\frac{1}{T_2} = \beta \cdot (\Delta x \cdot (1 - Y) \cdot \varepsilon)^2 + b'_0 \quad \text{small vessels}
\]

SE-based fMRI: contribution of small vessels increases with B_0

\[
\Delta R_2 = \frac{1}{\Delta T_2}
\]

Each curve corresponds to a different susceptibility difference between vascular and extra-vascular compartments (e.g. different contrast agents). SE-based fMRI shows a greater micro-vascular selectivity in the extra-vascular BOLD signal, because non-specific extra-vascular signal changes in the vicinity of large vessels are rephased.

IV = intra-vascular, EV = extra-vascular, \( \Delta x \) = magnetic susceptibility difference, \((1-Y)\) = deoxygenation of the blood, \( b_l, b_s \) = blood-volume fractions of small and large vessels, \( \alpha, \beta, \gamma \) constants \((\gamma = 1 \text{ free diffusion, } \gamma = 0.5 \text{ restricted diffusion})\). Susceptibility plots were taken from [20], and equations taken from [1].

First, the extra- and intravascular contributions to the BOLD signal change with field strength. While the extra-vascular signal contribution from small vessels increases with field strength, the intra-vascular contribution generally decreases as \( T_2 \) (blood) decreases with the field strength. Hence, SE-based functional contrast mechanism along with a long echo-time\(^3\) and high field strength provides for a high intrinsic spatial specificity of the functional signal. Second, unspecific signal contributions from downstream activation in large venous vessels can be separated by using bipolar gradients that dephase the intra-vascular signal or, third,

\(^3\) in order to ensure that the signal from blood water is completely dephased by the time the image is acquired.
venograms can be used in order to elucidate where large vessels are located. Furthermore, magnetization transfer selectively suppresses the signal that arises from water molecules in the vicinity of macromolecules. It therefore attenuates signal from tissue protons, because brain tissue contains macromolecules in high concentration or protons within capillaries due to their rapid exchange with tissue, but leaves protons in large vessels unaffected and can hence be used complimentary to diffusion-weighting [21].

**AUDITORY PERCEPTION AND ACOUSTIC NOISE**

Hearing is one of the most important human senses since it establishes, in combination with language, an invaluable source for social communication. The use of fMRI in the fields of auditory perception and language comprehension research however has been limited by the substantial acoustic noise generated by the MR scanner itself. Generally, the perturbations due to acoustic noise encompass a variety of physio-acoustic effects such as adaption (reduced BOLD response after prolonged exposure), loss of attention and temporary threshold shift (post-stimulation fatigue). The latter is mediated in part by a stapedial muscle reflex, which alters the perceived sound level and spectral characteristics of the presented stimulus, and hence causes acoustic stimuli such as speech and syllables experienced as relatively flat sounds.

In order to provide an overview about the impact and origin of ambient scanner noise, this chapter gives an introduction to the physiological underpinnings of auditory perception, the confounding effects of acoustic noise and available noise reduction strategies.

*Confounding Effects of Scanner Acoustic Noise*

The scanner acoustic noise not only produces activation in brain regions involved in auditory processing (direct confounding), but also interferes with the stimulus presentation itself when their spectral components overlap (indirect confounding).

Specifically, acoustic scanner noise evokes an elevated baseline signal, which in turn decreases the amplitude of the BOLD signal change and more generally, might even induce neuronal and/or hemodynamic saturation. Hence the functional response to the acoustic
stimulus does not simply add to the functional response evoked by continuous scanner noise, and the 'pure' functional response to the acoustic stimulus might not be extractable. Moreover, ambient scanner noise might also affect auditory perception by attention (as observed in an increased response in attention-related brain areas) and distraction (as measured by a drop in response in auditory brain areas). Furthermore, the extraction and recognition of an auditory stimulus in an obscuring acoustic background is hampered both at a cochlear level by forward masking (i.e. an overlap of spectral components) and at a cortical level by foreground-background decomposition effects (i.e. a larger response for low-intensity tones in background noise due to attentional effort). It has been shown previously that these effects account for a close relationship between intensity of scanner noise, behavioural performance, reaction time and BOLD signal change.

**Acoustic Noise Reduction in the MR Environment**

*Noise Sources & Pathways.* The acoustic noise generated by an MR scanner can be as high as 105-133 dB(A). Its primary source are the vibrational movements that are evoked by Lorentz forces that act upon the gradient coils. A conductor element carrying a current \( I \) placed into a magnetic uniform field \( B \) will experience a Lorentz force \( F_L = -B \times I \). The displacement of the gradient coils is dependent on the main magnetic field strength, the voltage amplitude and frequency waveform of the gradient switching (~ 1 kHz for a typical EPI sequence). The resulting vibration can be transmitted directly via air or through the gradient assembly when mounted on the patient tube (mechanical pathway).

A secondary source of acoustic noise are eddy-current-induced vibrations of metallic structures such as the cryostat inner bore and RF (radio frequency) coils. Other sources are ambient background noise from ventilation and cryogen reclamation systems (~70dB) or related to the RF pulses. The latter, RF "hearing", is mediated by the absorption of RF energy, which induces minute temperature elevation in head tissue that is subsequently followed by a pressure wave - which in turn is sensed by the hair cells of the cochlea via bone conduction.

A comprehensive summery of current strategies developed to reduce acoustic noise generation, transmission and interference with auditory stimulus presentation is shown in Figure 10.
**Hardware modifications.** Scanner acoustic noise can be substantially reduced by sealing the gradient assembly in a vacuum enclosure to block air-borne vibration propagation, and by supporting the gradient coil independently to block structure-borne vibration propagation. Moreover, using rubber damping materials and cushions or mounting it directly to the floor for immediate absorption of vibrational energy, the mobility of the gradient assembly can efficiently be restricted. While these modifications essentially block pathways carrying the noise and minimize sound propagation across components of the scanner, the construction of low-eddy-current RF coils and non-conducting inner bore might reduce an actual source of acoustic noise. Likewise, gradient coil arrangements can be found such that opposite Lorentz forces can mechanically be coupled by embedding the coil into stiff, non-compressible enclosures in a way that opposite forces in the gradient structure will annihilate (Lorentz force balancing, [27]). This design however, turns out to be problematic as visco-elastic properties of conductors might lead compression waves through the material and hence induce phase errors in the coupling of the coils.

**Active noise control systems** attenuate quasi-periodic noise effectively up to 1 Hz by injecting anti-phase noise into the noise propagation path. The performance, however, is dependent on the stability and numerical integrity of the algorithm. Notably the feedback loop might not be
able to cancel random noise since it might not be fast enough to detect and modify it, and might also perform less effective with faster imaging techniques.

**Soft gradient pulses.** It has been previously shown that the gradient waveform and the acoustic noise spectrum of an MR unit can be characterized as a linear system which can be described by an acoustic transfer function, or frequency response function [28]. Hence, the properties of the acoustic transfer function can be used in order to redesign the gradient waveforms, such that they do not contain frequencies for which the amplitude of the frequency response function is high (acoustic resonances), or that higher harmonics are suppressed by using low-pass filtered gradient pulses. It furthermore follows, that the maximum slope duration should be as long as possible and the number of slopes should be reduced, e.g. by merging gradient pulses – however at the expense of increased flow and motion sensitivity. However, these schemes are less effective in case of EPI/FLASH sequences, where the fundamental frequency is already in the audible range, and higher harmonics are less influential.

**Passive noise control,** such as by earplugs, headphones or insulating foam mattresses, which reduce vibrational coupling and bone conduction of acoustic noise, have shown to reduce acoustic noise by 10-30 dB(A). However, they might hamper verbal communication and may attenuate noise non-uniformly (i.e. perform poor at low frequencies, which is problematic for tonotopic mapping).

**FMRI acquisition strategies.** The majority of hardware-related noise control solutions limit the benefit of high performance gradient capabilities and are only capable of reducing but not eliminating ambient scanner noise completely. Hence, fMRI acquisition strategies have been adopted, which allow for long silent intervals, which are not contaminated by scanner acoustic noise. The specific design depends on as to whether silence is preferentially required during stimulus presentation (eliminate competition between the auditory stimulus at issue and the acoustic scanner noise), task performance (eliminate influence on the processing of the stimulus by attentional and distructional effects) or whether decoupling of the response (disentangle the functional responses evoked by the auditory stimulus and the acoustic scanner noise during the analysis) is required. Figure 11 depicts examples of a partially silent and a silent acquisition scheme. While in the former scheme data acquisition is disrupted only for a silent period during which the stimulus is delivered, the latter scheme exploits the
inherent delay between the onset of the stimulation and the succeeding functional response in order to separate the functional responses evoked by the scanner acoustic noise and the auditory stimulus at issue. Data are sampled sparsely in order to allow sufficient time for the functional response to the scanner noise to cease prior to the next trial. Compared to continuous scanning or the partially silent acquisition scheme, the 'silent' requires longer total imaging time in order to collect sufficient data required for comprehensive statistical analyses.

Figure 11: Silent and partially silent acquisition design for event-related (ER) functional MRI in human audition. Images are adopted from [29].
REFERENCES


SPATIAL RESOLUTION ENHANCEMENT USING SENSITIVITY-ENCODED ECHO-PLANAR IMAGING AT 3 TESLA IN A TYPICAL MOTOR PARADIGM

Conny F. Schmidt, Klaas. P. Pruessmann, Peter Boesiger

submitted to MAGMA

ABSTRACT

We employ a single-shot Sensitivity-Encoded (SENSE) gradient-echo EPI acquisition in order to enhance spatial resolution in a typical motor fMRI experiment at 3 Tesla. Using a commercial eight-channel receiver array and a SENSE acceleration factor of 2.7, functional time series were acquired with an acquisition matrix size of 192 x 192 within a readout time of 82 ms, yielding an effective in-plane resolution of 0.94 x 0.94 mm². When compared to a conventional acquisition which allows only moderate spatial resolution of (112 x 112 acquisition matrix), partial volume effects are considerably reduced in the high-resolution SENSE data. Consequently, clusters of activation were smaller, but delineated with fine structured detail, and underlying functional signal changes were larger than in the conventional data. Since SNR was reduced along with the voxel volume in the high-resolution SENSE data, statistical significance of clusters of activation was diminished, as apparent from significantly smaller z values. Our data suggest that fMRI can readily be performed with a spatial resolution adapted to detailed cortical functional topography. However, in all potential applications the specific behavior of spatial specificity and statistical sensitivity needs to be taken into account.
INTRODUCTION

Functional magnetic resonance imaging (fMRI) allows the non-invasive visualization of neuronal activity based on subtle changes in blood oxygenation during activation. The functional contrast has been termed blood oxygenation level dependent (BOLD) contrast [1], and is typically measured with $T_2^*$-weighted gradient-echo sequences using single-shot echo-planar imaging (sshEPI). The spatial resolution that can be achieved with these sequences is mainly limited by the number of phase encoding steps, that can be performed within a reasonable temporal resolution, broadening of the point spread function (PSF) due to $T_2^*$ decay and susceptibility effects. In the context of BOLD fMRI, additional constraints are the prescribed echo time and contrast-to-noise ratio (CNR).

Both $T_2^*$ blurring and susceptibility effects depend crucially on the encoding speed of the EPI acquisition trajectory. $T_2^*$ decay essentially acts as a low pass filter in k-space and thereby ultimately limits spatial resolution. Therefore the inherent spatial resolution limit is also determined by the acquisition time and thus by the acquisition speed in k-space, especially in phase-encoding direction in which the time between sampling points is large (proportional to the acquisition time for one read-out profile). Moreover, the frequency offset induced by magnetic susceptibility variation causes artifactual pixel shifts resulting in image distortion and signal dephasing, which likewise scale inversely with k-space speed. Hence it is crucial to maximize acquisition speed by fast switching of strong gradient fields. The gradient performance, however, is ultimately limited both technically and biologically by the risk of peripheral nerve stimulation.

One option to overcome these limitations is the transition to multiple-shot EPI or FLASH sequences [2-7]. With these approaches the problems of $T_2^*$ decay and field inhomogeneity can effectively be addressed. However, they require longer overall acquisition time and are prone to shot-to-shot inconsistencies, e.g., due to subject motion. Unlike these approaches, parallel imaging permits reducing the readout train by enhancing overall encoding speed [8-12], hence enabling high spatial resolution while preserving temporal resolution and robustness of a single-shot acquisition. In the present work this approach is pursued. The Sensitivity Encoding technique (SENSE, [13]) is employed in order to enhance the spatial resolution of a $T_2^*$-weighted gradient-echo sshEPI sequence for BOLD fMRI at 3 Tesla in a typical motor paradigm. High-resolution SENSE-sshEPI was then directly compared to conventional sshEPI acquisition.
While enhanced spatial resolution promises probing functional topography and neuronal network organization with superior spatial detail and specificity, smaller voxel volumes can equally benefit the sensitivity to functional signal changes by way of reducing partial volume effects with unactivated brain tissue [4]. However, the benefit is partly counteracted by signal-to-noise-ratio (SNR) drawback that results from the reduced voxel volume. Both these effects will cause differences in size and significance of clusters of activation when compared to conventional sshEPI acquisition at moderate spatial resolution. We study these effects by means of dedicated post-processing, such as spatial filtering and statistical threshold matching. Spatial filtering of the high-resolution data allows to perform a gradual trade-off between spatial resolution and SNR, while adaptive threshold matching (e.g. [14, 15]) accounts for the different degrees of "statistical sensitivity" in the data due to different underlying SNR.

High-resolution SENSE-sshEPI and conventional low-resolution sshEPI data are from here on referred to as HR and LR data, respectively. Smoothed with a 1.6 mm Gaussian kernel, the HR data are termed SHR.

MATERIALS AND METHODS

Subjects: Six healthy volunteers (three male, three female, age 29 ± SD 7 years), five right-handed (14.4 ± 1.7 scores according to Chapman & Chapman questionnaire [16]) and one left-handed (35 scores), participated in the study. Written informed consent was obtained from all volunteers prior to the examination. The study was in accordance with institutional ethical guidelines.

Stimulus and Paradigm: Stimulation consisted of bilateral opposite finger tapping alternating with rest in a block design of four 20 s epochs, which were signaled visually. A total of 80 volumes was acquired in 2 min 40 s runs. Subjects were trained prior to the examination to perform the task as regularly as possible with one tapping cycle per TR.

Data Acquisition: Measurements were performed on a Philips Intera 3 T whole body MR unit (Philips Medical Systems, Best, The Netherlands) equipped with a transmit-receive body coil and a commercial eight-element head receiver array (MRI Devices Corporation, Waukesha
WI, USA). Functional data were obtained from 6 transverse and coronal slices covering motor cortex with a spatial resolution of 0.94 x 0.94 x 5 mm³ (acquisition matrix 192 x 192) using SENSE-sshEPI with an acceleration factor of R = 2.7 and a partial Fourier acquisition of 77 % in the phase encoding direction. Other scan parameters were FOV = 180 mm, TR = 2000 ms, TE = 35 ms and θ = 75°. The experiment was repeated with conventional sshEPI, permitting only a resolution of 1.6 x 1.6 x 5 mm³ (matrix size 112 x 112) at otherwise identical scan parameters. Sampled with a pixel bandwidth in phase-encode direction of bw_{LR} = bw_{HR} = 12.3 Hz, the acquisition duration in both experiments was T_{acq} = 81.4 ms, and pixel bandwidths in readout direction were bw_{LR} = 1432 Hz and bw_{HR} = 836 Hz, respectively. Along with the larger acquisition bandwidth in phase-encode direction, sensitivity to geometric distortions is reduced from 130 µm/Hz in the LR acquisition to 77 µm/Hz in the HR acquisition. Furthermore, using smaller voxel volumes reduces intra-voxel dephasing, thereby preserving MR signal in typical "dropout" regions. Although these effects were not prominent at the selected slice locations in the present data, they will be important in regions of the brain that are prone to susceptibility effects, such as around sinuses and air-filled cavities.

Postprocessing and Analysis: Data post-processing and statistical analysis were carried out using FEAT software (FMRI Expert Analysis Tool) Version 5.00, which is part of FSL (FMRIB’s Software Library [17]). Functional time series were motion corrected using MCFLIRT, an automated linear (affine) registration tool [18]; brain/non-brain segmented using BET, which applies a surface model approach with a set of locally adaptive model forces [19], mean-based intensity normalized, and high-pass temporal filtered (using Gaussian-weighted LSF straight line fitting, with sigma = 30 s). The HR data were then treated in two different ways: first, no spatial filtering was applied to the data to allow direct comparison of spatial resolution. Secondly, the HR images were convolved with a Gaussian kernel of FWHM 1.6 mm. The kernel width was chosen such as to reconcile the FWHM of the point-spread function (PSF) of the LR conventional sshEPI data. The pixel-wise noise level was determined from the standard deviation over time of the residuals from a linear convolution model [20, 21] representing "shot-to-shot" temporal variance of the signal. This measure includes signal fluctuations that are not driven by stimulus-evoked activation, and accounts for thermal noise, physiologic processes, artifacts, and system drifts. SNR was assessed voxel-wise as the ratio of the mean signal and the noise level over time. Likewise,
the contrast-to-noise ratio (CNR) was computed as the ratio of the parameter estimates (beta values) of the GLM and the noise level, and averaged over significant clusters of activation.

**Theoretical SNR ratios:** Taking into account only image SNR, Weiger et al. [22] showed the SNR ratio between HR and LR data can be calculated as:

$$\frac{SNR_{HR-SENSE}}{SNR_{LR-Full}} = \frac{1}{g} \frac{V_{HR-SENSE}}{V_{LR-Flat}} \cdot \frac{\sqrt{(N_{PE} \cdot N_{FE})_{HR-SENSE}}}{\sqrt{(N_{PE} \cdot N_{FE})_{LR-Flat}}} \cdot \frac{\sqrt{bw_{LR-Flat}}}{\sqrt{bw_{HR-SENSE}}}$$  \[1\]

where \( V \) denotes the voxel volume and \( N_{PE} \) and \( N_{FE} \) the number of phase- and frequency-encoding steps, respectively. The acquisition bandwidth (\( bw \)) is proportional to the linear speed in k-space, thus

$$bw \propto \frac{N_{PE} \cdot N_{FE}}{T_{acq}}$$  \[2\]

The acquisition time \( T_{acq} \) was equal for both sequences, hence

$$\frac{SNR_{HR-SENSE}}{SNR_{LR-Flat}} = \frac{1}{g} \frac{V_{HR-SENSE}}{V_{LR-Flat}} \cdot \frac{1}{g} \frac{112 \times 112}{192 \times 192} = 0.34$$  \[3\]

The geometry factor \( g \) is ideally equal to 1 and somewhat larger with moderate acceleration. Hence, assuming thermal noise only and an ideal case of \( g = 1 \), the value of 34 % would represent an upper bound for the relative SNR underlying the HR as compared to the LR data.

**Statistical Analysis:** Time series statistical analysis was carried out using FILM (FMRIB's Improved Linear Model), a first-level General Linear Model analysis with time-series pre-whitening and local autocorrelation correction [23]. Z (Gaussianised T/F) statistic images were thresholded using clusters of activation determined by \( z > 3.5 \) and a (corrected) cluster significance threshold of \( p = 0.01 \) [24-26]. Additionally, the statistical threshold was adjusted for different CNR in two ways, first by raising the threshold for the low-resolution data and second by lowering the threshold for the high-resolution data. In both cases the adjustment criterion was to match the product of the statistical threshold and the CNR. According to this criterion the LR data were thresholded either with \( z > 4.3 \), or the HR data with \( z > 2.8 \), respectively.
RESULTS

Signal to Noise Ratio

HR EPI images taken from the functional time series clearly delineate fine anatomical detail (Figure 1a). Corresponding SNR maps computed from the LR, HR and SHR data clearly reveal reduced SNR in the HR acquisition, which can partly be recovered by appropriate spatial filtering (Figure 1b). Importantly, spatially varying noise amplification along the phase encode direction (AP in the transverse and RL in the coronal slices) is readily appreciable in SNR reduction in the center of the SNR maps and is characteristic to parallel MRI. It results from non-unitary operations during image reconstruction and can be quantified by the so-called g factor [13].

SNR in the HR data is reduced to 44 % as compared to the SNR in the LR data (Tables 1 & 2). Theoretical SNR calculation (Eq. [3]) however predicts a SNR reduction to 34 %. Therefore our data indicate, that SNR loss resulting from smaller voxel volume and parallel acquisition in the HR data was less than predicted from theory, which, however, accounts only for image noise. Primarily a consequence of reduced voxel volume, appropriate spatial filtering on the HR data almost (86.1 % in the transverse and 89.4 % in the coronal acquisition) reproduced the same SNR as obtained with the LR acquisition.

Functional Activation Maps

We identified significant clusters of activation bilaterally in sensorimotor neocortex and supplementary motor area during bilateral opposite finger tapping (Figures 2 & 3). The HR acquisition uncovers more detailed and fine structured clusters of activation. In particular, a narrow fissure that does not show functional activation and pervades a region with highly significant activation, designates a persistent feature in the pattern of activation in the HR, SHR and HR (z > 2.8) data, but is not discernible in the LR data (marked by arrows in Figures 2 & 3).
Figure 1: **a.** Sample transverse (top row) and coronal EPI image (bottom row) taken from the functional time series. **b.** Representative SNR maps, computed from the LR, HR and SHR time series for the same slices as in **a.** The maps are scaled to the same maximum to facilitate comparison. SNR of HR images is remarkably reduced as compared to the SNR in the LR images. SNR can partly be recovered by appropriate spatial filtering, as shown in the SHR SNR images. The SNR drop in the image centre results from noise amplification along the anterior/posterior direction (transverse slice acquisition) and left/right direction (coronal slice acquisition) respectively, in which SENSE acceleration was applied.
Figure 2: Statistical (z) and functional (PE) activation maps. Top row: Shown from the left to the right are statistical parametric maps (z maps) revealed from and overlaid on a typical transverse section through one subject's motor cortex, illustrating activation in primary motor cortex and supplementary motor area as revealed from LR, HR and SHR data (p < 0.01, z > 3.5 and 2.8 respectively). Middle row: Enlarged detail. Bottom row: Parameter estimate (PE) maps provide a measure for the strength of the functional response, i.e. highlight underlying physiological changes. Remarkably, a small fissure of almost no activation separates two highly activated areas as marked by arrows. Partial volume effect is apparent in the region marked by a black rectangle, showing a large functional response at high spatial resolution, but only moderate functional response at low spatial resolution (LR and SHR). L = R, R = L.
In order to highlight the impact of partial volume effects in the LR as compared to the HR data, we selected a representative section of a cluster of activation in Figure 2. While a larger and more significant (larger z values) cluster of activation was revealed in the LR data, the underlying functional response (as quantified by the PF) was comparatively low. Vice versa, the same cluster exhibits a large functional response at HR, which however transforms only into moderate z values as a result of lower SNR. These data suggest that the functional response is diminished by partial voluming with adjacent non-activated brain tissue in the LR data, however, despite the large underlying functional effect, the statistical sensitivity in the HR data is reduced by lower SNR. In order to further highlight the significance of partial volume effects, we chose a simultaneous plot of BOLD signal time courses for adjacent

**Figure 3:** Statistical (z) and functional (PF) activation maps, as in Figure 2, revealed from and overlaid on a typical coronal section through motor cortex.
voxels placed within a highly structured cluster of activation (Figure 4). As readily appreciated from the signal time courses, functional signal change was largest in the HR. However, as a result from reduced SNR concurrent with smaller voxel volumes, signal fluctuations not driven by the stimulus were as well markedly larger. Both were diminished by spatial filtering in the SHR and likewise in the LR data, accompanied by enhanced statistical significance (z values). Importantly, the small fissure identified in the HR data (also marked by arrows in Figures 2 & 3) indeed qualified as region that does not show stimulus-driven variation in the functional time series, while a functional response was clearly revealed in adjacent voxels. In the SHR and LR data, this fissure does not appear, because the temporal functional signal fluctuation from these adjacent voxels contributed markedly to its own signal time course – pinpointing to a partial volume effect from low spatial resolution.

Notably, the size of significant clusters of activation is significantly smaller in HR than in LR data (p < 0.04, 58.0 ± 6.1 % for transverse and 47.3 ± 1.6 % for coronal acquisition). Clearly, this effect is inherent to the specific traits of the HR as compared to the LR acquisition. Subject motion did not exceed one third of the voxel dimensions and therefore can not account for this effect. Those regions that did not exhibit significant activation in the HR but showed a significant response in the LR data, were almost exclusively located in regions were the LR acquisition revealed activation with low statistical significance.

Yet at the expense of fine structured detail, spatial resolution can be traded with appropriate spatial filtering in the SHR data in order to enhance SNR and therefore to invigorate the functional signal against the underlying noise. Cluster size in the SHR data then reached 117.8 ± 10.7 % (transverse) and 94.3 ± 3.2 % (coronal) of the average cluster size obtained in the LR data.

Average z values within significant clusters of activation differed only slightly between the HR and LR data and were almost identical in the SHR and LR data (Tables 1 & 2). The distribution of z values shows a similar pattern for the HR and LR data despite different distribution heights, which reflects reduced overall activated brain volume at HR.
Figure 4: Functional signal time courses (ΔBOLD BOLD) for LR, HR and SHR data obtained from regions highlighted by arrows in Figures 2 & 3, which are also marked by a red square in the detail section in the lower left corner. Functional signal time courses are overlaid on statistical parametric maps, allowing direct inference between the underlying physiological signal change and the statistical effect. Partial volume effects can directly be assessed from voxel-wise comparison.
As an immediate measure of statistical sensitivity, we determined the CNR in LR, HR and SHR data (Tables 1 & 2). The CNR ratio between the HR and LR data was 76.5 ± 2.3 % (transverse) and 85.9 ± 3.3 % (coronal). Based on these ratios the statistical threshold was adapted to $z > 2.8$ in the HR data and alternatively raised to $z > 4.3$ in the LR data, in order to match for the different statistical sensitivity underlying the HR and LR time series. The HR data with a statistical threshold $z > 2.8$, however, still yielded only smaller clusters of activation as compared to the LR data (86.0 ± 7.2 % (transverse) and 74.7 ± 1.8 % (coronal)). Vice versa, raising the statistical threshold of the LR data to $z > 4.3$, the relative cluster size as obtained with the HR was comparatively smaller (76.7 ± 6.6 % (transverse) and 68.4 ± 3.1 % (coronal)). Unlike spatial filtering, adjusting the statistical threshold conserved the gain in information attained by the higher spatial resolution, though compromising on statistical significance.

Average PEs within significant clusters of activation were comparatively higher in the HR than in the LR data ($p < 0.001$), however are not significantly lower in the SHR than LR data (Tables 1 & 2). Particularly, the underlying PE distributions in the HR and LR data are nearly identical for PEs > 5 (Figure 5). Notably, the distribution peak for the HR data is located at 3.8, while the LR PE distribution peaks at 2.0. Moreover, virtually no significant voxel has been revealed below PEs of 2.0 in the HR data, indicating that the difference in cluster size observed between HR and LR data evolved primarily from regions with low underlying functional signal. The SHR distribution essentially, and expectedly, mimics the characteristics of the LR PE distribution. Adjusting the statistical threshold for the LR data to $z > 4.3$, in order to account for the higher SNR, reduced the distribution high at low PEs. The distribution peak was shifted to higher PE of 2.5, however the distribution still showed more voxels at low PEs as compared to the HR distribution. Likewise, diminishing the statistical threshold for the HR data to $z > 2.8$ enhanced the number of voxels at small PEs, but still not revealed voxels for PEs well below 2.0.

These results suggests, that clusters of activation might have been blown up by partial voluming in the LR data. Furthermore, activation in regions with underlying low functional signal changes is not necessarily "lost" in HR data, but rather masked by reduced SNR.
### Table 1: Measures of statistical sensitivity and functional activation (transverse slice acquisition).

<table>
<thead>
<tr>
<th></th>
<th>SNR</th>
<th>CNR (act)</th>
<th>cluster size [10³ mm³]</th>
<th>z</th>
<th>PE</th>
</tr>
</thead>
<tbody>
<tr>
<td>HR</td>
<td>30.3 ± 0.6</td>
<td>1.7 ± 0.1</td>
<td>17.0 ± 4.4</td>
<td>5.4 ± 0.2</td>
<td>6.3 ± 0.4</td>
</tr>
<tr>
<td>LR</td>
<td>68.4 ± 2.2</td>
<td>2.2 ± 0.1</td>
<td>30.9 ± 9.1</td>
<td>5.9 ± 0.2</td>
<td>4.1 ± 0.4</td>
</tr>
<tr>
<td>SHR</td>
<td>61.1 ± 1.5</td>
<td>2.0 ± 0.1</td>
<td>34.3 ± 8.7</td>
<td>5.7 ± 0.2</td>
<td>3.7 ± 0.3</td>
</tr>
<tr>
<td>HR/LR [%]</td>
<td>44.4 ± 0.8</td>
<td>76.5 ± 2.3</td>
<td>58.0 ± 6.1</td>
<td>90.8 ± 1.5</td>
<td>155.7 ± 9.8</td>
</tr>
<tr>
<td>SHR/LR [%]</td>
<td>89.4 ± 1.5</td>
<td>87.9 ± 1.9</td>
<td>117.8 ± 10.7</td>
<td>96.7 ± 1.3</td>
<td>91.4 ± 4.9</td>
</tr>
</tbody>
</table>

SNR and CNR, cluster size, mean z values and parameter estimates as obtained in the HR, LR and SHR data. SNR was averaged across the acquired brain volume, while CNR, z values and PEs were averaged across significant clusters of activation. Given are absolute values averaged over subjects (mean ± SEM) and the ratios HR/LR and SHR/LR data, respectively. Ratios were determined for each subject individually and then averaged across all subjects.

### Table 2: Measures of statistical sensitivity and functional activation (coronal slice acquisition).

<table>
<thead>
<tr>
<th></th>
<th>SNR</th>
<th>CNR (act)</th>
<th>cluster size [10³ mm³]</th>
<th>z</th>
<th>PE</th>
</tr>
</thead>
<tbody>
<tr>
<td>HR</td>
<td>28.8 ± 0.7</td>
<td>1.7 ± 0.1</td>
<td>14.1 ± 3.7</td>
<td>5.1 ± 0.2</td>
<td>5.9 ± 0.3</td>
</tr>
<tr>
<td>LR</td>
<td>65.2 ± 1.3</td>
<td>1.9 ± 0.1</td>
<td>29.1 ± 7.2</td>
<td>5.6 ± 0.2</td>
<td>3.7 ± 0.3</td>
</tr>
<tr>
<td>SHR</td>
<td>56.1 ± 1.3</td>
<td>1.9 ± 0.1</td>
<td>27.3 ± 6.6</td>
<td>5.5 ± 0.2</td>
<td>3.5 ± 0.2</td>
</tr>
<tr>
<td>HR/LR [%]</td>
<td>44.2 ± 1.0</td>
<td>85.9 ± 3.3</td>
<td>47.3 ± 1.6</td>
<td>93.2 ± 3.7</td>
<td>159.4 ± 5.2</td>
</tr>
<tr>
<td>SHR/LR [%]</td>
<td>86.1 ± 1.2</td>
<td>97.8 ± 3.2</td>
<td>94.3 ± 3.2</td>
<td>98.8 ± 3.7</td>
<td>93.7 ± 2.9</td>
</tr>
</tbody>
</table>

Summarized are SNR and CNR, and cluster size, mean z values and parameter estimates as obtained in the HR, LR and SHR data. Same notation as in Table 1.
Figure 5: Comparison of distributions of z values and PEs for LR (grey), HR (black) and SHR data (grey, thin line) for transverse (a.) and coronal (b.) slice acquisition. Size of activated brain volume is directly reflected in the area under the distribution curve. The PE distribution for PEs > 5 merge for HR and LR data, while voxel with low PEs (< 2.0) are hardly detected in the HR data. Raising the statistical threshold for the LR data (z > 4.3) essentially damps the PE distribution height at low PEs (grey, dotted line). However the PE distribution in the LR data still shows a larger number of voxels with low PEs as compared to the PE distribution in the HR data. Lowering the statistical threshold for the HR data (z > 2.8) reveals a larger number of voxels with low PEs, but does not show voxels for PEs well below 2.0 (black, dotted line), which indicates partial volume effect.

DISCUSSION

In the present study we exploited the enhanced acquisition speed of Sensitivity Encoding in order to boost spatial resolution in a typical motor experiment. While preserving a temporal resolution of 2 s, the spatial resolution of a standard gradient-echo sshEPI sequence was increased by a factor of 1.8 in both in-plane image dimensions, and thus by a factor larger than 3 in terms of voxel size. As a consequence of reduced partial voluming, clusters of activation in the HR data were smaller and delineated with fine structured detail. Likewise the underlying functional response was larger in the HR than in the LR data. However, since SNR was comparatively reduced, statistical sensitivity was diminished in the HR data.
Statistical sensitivity to functional signal changes depends crucially on temporal signal stability. Therefore all temporal signal fluctuations were taken into account while calculating SNR values. Based on "image SNR" alone, the SNR ratio between HR and LR data revealed a SNR drawback in the HR data that was less than straightforward theory predicted. A similar effect was previously observed by Preibisch et al. [27] at a magnetic field strength of 1.5 Tesla. However, the theoretical SNR approach in Eq. [3] does not account for significant temporal autocorrelation of physiological noise [28], and hence is likely to overestimate the SNR drawback. Previous studies have demonstrated that the physiological noise increases with the underlying signal and hence limits the maximum disposable SNR [9]. These findings therefore suggest that our LR data might be more detrimentally affected by physiological noise than the HR data and thus might have revealed comparatively lower SNR than what was predicted from pure thermal noise. Notably, the importance of (signal-dependent) physiological noise on time course signal stability has recently been investigated in a fMRI study performed at 7T [29]. These findings implied that spatial smoothing of high-resolution functional images to the desired (low) spatial resolution might be more advantageous rather than a low resolution acquisition in terms of physiological noise contribution [30]. One other likely reason for the higher than expected SNR in the HR as compared to the LR data might be a comparatively reduced $T_2^*$ decay at low spatial frequencies in the SENSE acquisition.

The CNR as a measure of statistical sensitivity is even more difficult to analyze as it also depends on the size and shape of the functional pattern of activation. An increasing body of studies has aimed to optimize the voxel size for fMRI in terms of CNR, suggesting optima between 3.4 - 8 mm$^3$ [31] and 16 mm$^3$ [32]. With voxel volumes of 12.8 mm$^3$ and 4.4 mm$^3$ both LR and HR data in the present work fall within this range.

**Functional Activation Maps**

Statistical parametric maps derived from the HR data clearly revealed more spatial structure. Partial voluming was markedly diminished when compared the LR data. While HR acquisition allows to localize activation with high spatial specificity (e.g.[33, 34]), statistical sensitivity is reduced along with lower SNR, as apparent from lower z values. This effect will particularly be important in regions with low underlying functional signal change, since these might not be detected with sufficient statistical confidence (sub-threshold voxels). We studied
this effect by way of a simultaneous plot of functional signal time courses in adjacent voxels within a highly structured cluster of activation (Figure 4) as well as dedicated post-processing, such as spatial filtering and adaptive statistical thresholding.

Regions with inherently small functional signal change: In the close vicinity of non-activated brain tissue, small functional signal changes might not be localized with an LR acquisition due to partial voluming, which decreases the measured functional signal change. While partial volume effects are effectively reduced at high spatial resolution, our data show the benefit of spatial resolution partially hidden by reduced SNR (and hence lower statistical sensitivity). In this case, statistical threshold adoption is an option for localizing such activation. As shown in the present data (HR z > 2.8 data in Figures 2 & 3), significant clusters of activation with low underlying functional response can be uncovered in this fashion. Clearly, the interpretation of such results grows more difficult as the statistical confidence decreases and false positive activation might more likely occur. Detecting tiny regions with small functional signal changes thus remains challenging even with high resolution scans due to lacking statistical power.

Regions with inherently large functional signal changes: In regions with high underlying functional signal changes, a clear benefit from HR acquisition was demonstrated in our data, as clusters of activation were delineated with excellent structural detail. Exemplarily we selected a region of interest covering a small fissure of no activation located between two highly significant activation clusters. This fissure was only resolved in the HR data and persistent even at lower statistical significance threshold (HR, z > 2.8). Furthermore, since the functional signal time course obtained along this fissure did not showed a stimulus-evoked signal fluctuation, the interpretation that with LR this structure cannot be detected simply due to insufficient spatial resolution is strongly supported. Similarly, a tiny region with high underlying functional signal change (as marked by a black rectangle in Figure 2), exhibits high PEs in the HR data, while only moderate PEs were apparent at low spatial resolution (LR and SHR data), illustrating a partial volume effect with surrounding non-activated brain tissue, and larger cluster size in the LR originated simply from larger voxel volume used in the LR acquisition. Hence, the HR acquisition enhanced the precision with which functional boundaries that segregate patches of cortex with high underlying functional activation and non-activated brain tissue can be delineated.
**PE distributions:** The HR acquisition not only revealed more spatially detailed activation maps, but also showed significantly larger PEs than the LR data. Consistent with previous studies, we demonstrated the effect resulting from partial voluming [4]. However, we found the overall effect of partial voluming not the only cause for reduced size of clusters of activation in the HR as compared to the LR data, since reduced SNR diminished the statistical sensitivity of the HR data. Partial volume effects came apparent in the PE distributions. When "CNR matched" (by adaptive statistical threshold) HR and LR distributions were compared, the smaller number of voxels with PEs well below 2.5 in the HR $z > 2.8$ data as compared to the LR data, or likewise the larger number of voxels with PE < 2.0 in the LR $z > 4.3$ as compared to the HR data, clearly indicated partial voluming in the LR data. SHR data essentially mimicked the characteristics of LR distribution, as expected.

**Intrinsic spatial resolution of the functional signal:** Since spatial resolution is not only a matter of technical advances but also crucially depends on the spatial specificity of underlying functional contrast. Therefore, the intrinsic spatial resolution of the BOLD response itself needs to be critically addressed. An indirect measure of neuronal activity, the BOLD response is primarily driven by a vascular response and is thus smooth on a spatial and temporal scale. The spatial resolution limit of BOLD fMRI is hence determined by the vascular (venous) geometry – and the neuro-vascular coupling mechanism. As shown by Duvernoy et al. [35], a penetrating vein in cortical layers 3 and 4 drains a 0.75 - 4 mm$^3$ volume of cortical grey matter tissue. Analyzing the regular, self-similar topography of the venous bed, a relationship was previously deduced that links the spatial extent of the underlying activity to the effect of draining veins and therefore spatial specificity [36]. This model may not be strictly applicable to the data obtained in our work, because cluster size depends on the significance threshold and deviates from the assumed circular shape. Nevertheless, it is still useful to derive at least an estimate of the intrinsic resolution of the BOLD signal in the present case. In order to apply Eq. [11] from [36], the required model diameter can be set to the smaller dimension (approximately 4 pixels) of the main cluster of activation in the HR statistical map (Fig. 1), yielding an upper limit of 1 mm downstream activation ($L = 0.42 * \sqrt{6.15 \text{ mm}^2}$). This estimation suggests that the spatial resolution used in our study lies well within the spatial specificity of the BOLD signal. It should furthermore be mentioned that previous studies used the more focussed early negative BOLD response in order to localize neuronal activation with high spatial specificity [34, 37, 38]. Moreover, gradient-echo techniques are sensitive to static dephasing effects and thereby are inherently sensitive to both, macro- and microvasculature. In
a spin-echo experiment static dephasing effects are essentially refocused, and diffusion effects in the static (susceptibility) field gradients come more apparent. The change of the $T_2$ relaxation rate exhibits a maximum at the capillary radius range ($2 - 3 \mu m$) and decreases with increasing vessel size [1]. Consequently, spin-echo sequences are comparatively more sensitive to oxygenation changes in microvasculature (provided the intravascular signal contribution is eliminated) and thus prove a functional contrast mechanism more closely tailored to the actual site of activation [39, 40]. Importantly, all these techniques can be combined with parallel acquisition propelling spatial while preserving temporal resolution.

CONCLUSION

We have demonstrated in the present study that the speed benefit of parallel imaging permits BOLD fMRI with enhanced spatial resolution. The spatial resolution of a conventional single-shot gradient-echo EPI sequence was enhanced by a factor of 1.8 in both in-plane dimensions and more than threefold in terms of voxel volume. Statistical parametric maps obtained from high resolution data delineated more spatially detailed and fine structured clusters of activation than conventional fMRI at moderate spatial resolution, especially in regions with large underlying functional response. However, the potential of high resolution fMRI was limited by statistical sensitivity in regions with low underlying functional activity.

Improved spatial resolution is crucial in many applications, particularly in studies focusing on details of cortical functional topography. However, in all potential applications the specific behavior of spatial specificity and statistical sensitivity needs to be taken into account.

Acknowledgments

The authors wish to thank for the continuing support of Philips Medical systems and the financial support by ETH Zurich (Strategic Excellence Project TH 7/02-02).
REFERENCES


17. www.fmrib.ox.ac.uk/fsl.


SENSITIVITY-ENCODED (SENSE) ECHO PLANAR FMRI AT 3T IN THE MEDIAL TEMPORAL LOBE

Conny F. Schmidt, Nadia Degonda, Roger Luechinger, Katharina Henke, Peter Boesiger


ABSTRACT

Sensitivity Encoded (SENSE) parallel imaging allows accelerating the acquisition process at preserved spatial resolution. The associated increase of the acquisition bandwidth is advantageous when susceptibility gradient fields shift the MR signal within k-space. Despite its potential benefits for reducing susceptibility-related image artefacts, the intrinsic signal-to-noise-ratio (SNR) deteriorates with the application of SENSE as a result from undersampling and from spatially varying noise amplification during the image reconstruction process. Hence the question emerged as to whether at all the SENSE technique can be advantageously applied for fMRI - especially in the medial temporal lobe, where noise amplification is most critical. We conducted two fMRI experiments and compared SENSE single-shot (ssh) echo-planar imaging (EPI) at acceleration factors of $R = 2.0, 2.4, 2.7, \text{ and } 3.0$ with conventional sshEPI at TE of 22 and 35 ms in a face-profession learning task. Susceptibility related image distortion was apparently reduced with SENSE as compared to conventional EPI, yet acceleration beyond 2 revealed only subtle further reduction of geometric distortions and enhanced encoding speed became at remarkable expense of statistical significance in the functional activation maps in the medial temporal lobe - except for an unexpected...
improvement at $R = 3.0$. Moreover, statistical power increased with a TE of 35 ms compared to 22 ms. Our data suggest that SENSE-sshEPI acquisition with $R = 2.0$ is suitable for fMRI experiments aiming at the detection of medial temporal activation.

INTRODUCTION

Functional magnetic resonance detects changes in blood oxygenation and is based upon changes in deoxyhemoglobin content, that in turn provokes a local change in the susceptibility field. The functional contrast is termed “blood oxygenation level dependent” (BOLD) contrast [1], and the functional signal is typically measured with gradient-echo sequences combined with echo planar imaging (EPI) acquisition. Gradient-echo sequences, however, are not exclusively sensitive to functionally related susceptibility differences, but are also prone to susceptibility gradients in the proximity of air-tissue interfaces and to other field inhomogeneities that cause off-resonance phase errors, which are more pronounced at higher field strengths.

Local susceptibility field gradients cause neighboring spins to precess at different frequencies. Since the MR signal is localized by frequency encoding [2], the presence of susceptibility gradients manifests in image distortions such as spatially varying shifts and a position dependent magnification of the object. Moreover, susceptibility gradients can shift the gradient echo within k-space leading to a change of the effective echo time, which in turn results in a heterogeneous BOLD sensitivity [3] across the image. In the extreme case, the signal is shifted outside the acquisition window, which results in complete signal loss, most apparently in orbitofrontal cortex and the medial inferior temporal lobes. Increased sampling bandwidth and shorter acquisition time, respectively, help reducing these artifacts simply by limiting the time during which these can evolve. Both can effectively be achieved with parallel imaging.

Parallel acquisition techniques such as Sensitivity Encoding (SENSE, [4]), make use of receiver coil arrays for image acquisition, which allow to reconstruct the full-field-of-view images from undersampled data by using information from spatially varying coil sensitivity profiles. In turn, the undersampling leads to shorter echo train length and faster k-space coverage. However, compared to full Fourier encoding, SENSE comes along with a signal-to-
noise ratio (SNR) penalty resulting from lower sampling density and spatially varying noise amplification from non-unitary operations in the reconstruction process, which depends on coil sensitivity distribution and is characterized by geometry factor maps (referred to as g-maps thereafter [4]).

Therefore, the benefits from reduced susceptibility artifact and drawbacks from reduced SNR need to be balanced and the application of parallel imaging has to be carefully adapted to the study-specific regions of interest. Initial studies already demonstrated the potential of parallel imaging for fMRI at 1.5 Tesla field strength in terms of artifact [5, 6] and gradient acoustic noise reduction [7], and for high spatial resolution fMRI [8, 9]. These studies compared robust fMRI activation in primary auditory cortex and motor cortex evoked by simple auditory and motor tasks. In most fMRI studies involving higher cognitive functions, however, regions of interest are more affected by susceptibility artifact and often face with less robust and weaker BOLD signal changes. In particular, the medial temporal lobe structures, which include the hippocampus and parahippocampal gyrus, are not only highly affected by susceptibility-related geometric distortions, but exhibit only modest BOLD signal changes of about 1% as compared to 2–3% typically found in primary sensory areas [10]. While with the contrast to noise benefit of high magnetic field strength of 3 Tesla functional sensitivity is increased, susceptibility artifacts become increasingly worse, requiring fMRI procedures specifically tailored to detect activation in the medial temporal lobe at 3 Tesla field strength.

While our parallel imaging approach focuses on reducing intra-voxel dephasing on a temporal scale, i.e. by shortening the acquisition time during which susceptibility effects can evolve, others aimed at reducing susceptibility inhomogeneities comprised within each image voxel by reducing slice thickness and voxel volumes, i.e. on a spatial scale [11-13]. Although the results from this approach are quite convincing, it prohibits whole brain acquisition within reasonable repetition times. Techniques relying on gradient compensation, on the other hand, can only correct for susceptibility inhomogeneities in specified local regions while affecting other regions oppositely [3]. Moreover, we additionally aimed at tracking another temporal aspect of intra-voxel dephasing namely reducing the echo time, which sets the total time scale for intravoxel dephasing. However, this goes along with a concurrent reduction of BOLD sensitivity (e.g. [14]).
We carried out two fMRI studies in order to compare different imaging parameters for fMRI in the medial temporal lobe at a field strength of 3 Tesla. In the first fMRI experiment, we compared image acquisition with and without SENSE at two different echo times, namely TE = 22 ms and 35 ms. This experiment was performed to investigate whether or not Sensitivity Encoding is suitable for fMRI in the medial temporal lobe. In the second fMRI experiment, we compared acquisition with SENSE acceleration factors of 2.0, 2.4, 2.7, and 3.0 in a new group of subjects. Herewith we aim investigating to what extent Sensitivity Encoding can be applied without critically compromising statistical significance resulting from increasing noise amplification due to image reconstruction process and undersampling in k-space. In both experiments, we used a paired-associate learning paradigm [15] in order to evoke activation in the medial temporal lobe (MTL) bilaterally. Learning success was assessed with a cued recall task applied immediately after the functional scan. Retrieval performance was a covariate in the statistical comparison between acquisition protocols to partial out variance underlying inter- and intra-individual cognitive differences. We also evaluated activation with respect to other regions of interest commonly activated during this task, such as the left and right inferior frontal gyrus (IFG) and the left and right fusiform gyrus (FG).

METHODS

Memory Tasks: We used a face-word learning paradigm (Figure 1), which reliably evokes a functional response in the left and right MTL [15]. Experimental stimuli consisted of pairs of a face [16] and a written profession, which alternated with head contours without physiognomy (baseline condition). The learning instruction in the experimental condition was to imagine each presented person in a scene of the indicated profession and to report by button press whether it was easy or hard to imagine such a scene. This imagination task automatically induces a semantic binding and encoding of the face and the profession. The instruction for the baseline task was to decide and indicate by button press whether the left or right ear of a head contour was larger. This baseline task minimally engages the medial temporal lobe because it captures participants’ attention with uniform stimulus material, and thereby prevents uncontrolled mentation and storage. All subjects practiced the tasks before scanning with training material. In all conditions stimuli were presented for 6 s, with four stimuli presented per block (blocks of 24 s) and three blocks per condition in each fMRI time-series. Each block was initialized by a 2 s instruction slide with “E” indicating the ear task
and "C" indicating "combine face and profession" for the imagery task. Scan time for each fMRI time series was 2 1/2 min. After each complete time series, faces were displayed again as memory cues (without professions), each for 6 s, with the instruction to recall the learned profession to indicate learning success. Stimuli were presented using an inhouse software (SCOPE V2.5.4 Display Program, Max R. Duersteler, University Hospital Zurich, Switzerland). Button responses were recorded with a fiber-optic response pad with a four inline button hand-held box (Current Designs, Inc., Philadelphia, USA).

Figure 1: Functional paradigm. Participants were instructed to learn the face-profession combinations by use of an imagery task. This task required subjects to imagine the presented person acting in a scene of the indicated profession, thus inducing a semantic association and encoding of the profession to the person. In the baseline condition, head contours were presented with the instruction to decide which ear was larger, left or right. Each stimulus was presented for 6 s with four stimuli per block and three blocks per condition. Each block began with a 2 s presentation of an introduction slide with "E" for "ear" and "C" for "combine face and profession". The whole fMRI time series lasted 2 1/2 min. Faces were reproduced from the book "Heads" (1985) by permission of Alex Kayser.

Subjects: 28 healthy volunteers participated in our fMRI studies, and were divided into two age- and sex-matched groups. The first group of subjects (11 males and 3 females, mean age 25 ± 4 years) participated in Experiment 1 (image acquisition with and without SENSE). The second group of subjects (12 males and 2 females, mean age: 26 ± 5 years) participated in Experiment 2 (comparing four SENSE acceleration factors). Written informed consent was
obtained prior to examination. One subject was excluded from analysis because of poor learning and retrieval performance (< 8 correctly recalled professions out of 12).

Data Acquisition: Measurements were performed on a Philips Intera 3T whole-body MR unit equipped with a transmit-receive body coil and a commercial eight-element head coil array (MRI Devices Corporation, Waukesha WI, USA). Functional data were obtained from 31 transverse slices covering the whole brain with a spatial resolution of 2.75 x 2.75 x 4 mm³ (acquisition matrix 80 x 80). Slices were oriented parallel to the ACPC line. In Experiment 1, conventional ssh-EPI data were compared to data acquired with SENSE-sshEPI with acceleration factor R = 2.0, at both TE = 22 ms and TE = 35 ms. The conventional full Fourier encoded ssh-EPI data at TE = 22 ms needed to be sampled with a partial Fourier factor of 82 % because the pre-echo period was too short for the applied scan matrix. The different acquisition conditions are from here on referred to as conv-te22, conv-te35, sense-te22 and sense-te35. In Experiment 2, a SENSE sshEPI was applied with different acceleration factors R = 2.0, 2.4, 2.7, and 3.0. The acquisition protocols used in Experiment 2 will be termed sense-2.0, sense-2.4, sense-2.7 and sense-3.0 thereafter. In both studies, the order of applied measurement protocols was systematically varied across participants to evenly distribute fMRI data variance induced by sequence effects such as habituation. Additional acquisition parameters were TR = 3000 ms, TE = 35 ms and θ = 82°. Coil sensitivity reference maps for sensitivity calibration used in the SENSE reconstruction were acquired with 3.75 x 3.75 x 4 mm³ spatial resolution, acquisition matrix 64 x 64, TE = 1.54 ms, TR = 7.0 ms, θ = 7°. A 3D T1-weighted scan was obtained for anatomical reference with 1 x 1 x 0.75 mm³ spatial resolution, acquisition matrix 224 x 224, TE = 2.30 ms, TR = 20 ms, θ = 20°.

Postprocessing and Analysis were performed using FEAT (FMRI Expert Analysis Tool) Version 5.1, which is part of FSL (FMRIB's Software Library, [17]). The following pre-statistics processing was applied to the functional data: motion correction using MCFLIRT, an automated linear (affine) registration tool [18]; brain/non-brain segmentation using BET, which applies a surface model approach with a set of locally adaptive model forces [19]; spatial smoothing using a Gaussian kernel of FWHM 5 mm; mean-based intensity normalization of all volumes by the same factor; and finally high-pass temporal filtering (Gaussian-weighted least-squares straight line fitting, with sigma = 50.0 s). Motion-related displacement during the time course for all subjects and conditions did not exceed one third of
the pixel dimensions (0.6 mm). It should be noted here, that spatial smoothing does not affect the conclusions of the study by potentially differential effects in connection with the SENSE reconstruction. Intrinsic noise correlation originating from the reconstruction process only occurs between pixels that collapse after Fourier transform of the single-coil data and hence are separated by multiples of the reduced FOV size [4]. Spatial smoothing, however, incorporates closely neighboring pixels and therefore interference between noise correlation and spatial smoothing can be ruled out. Yet, we determined signal-to-noise-ratios from both, unsmoothed and smoothed, motion corrected and high-pass filtered data, while statistical analysis was performed only on the spatially smoothed data.

The noise level was determined pixel-wise from the signal time course of the functional time series as the standard deviation of the residuals from General Linear Model (GLM) fit, i.e. signal fluctuations that cannot be modeled by the General Linear Model. This measure is sensitive to all time course variations caused by physiologic processes, artifacts or system drifts and therefore describes the "shot-to-shot" temporal variability of the signal. As an immediate measure of functional signal stability the noise levels were assembled in noise maps. Signal-to-noise-ratios (SNR) were determined on a pixel-by-pixel basis as the ratio of the mean signal over time and the noise level.

Statistical analysis of the functional time-series was performed using FILM (FMRIB's Improved Linear Model), a first-level General Linear Model analysis with time-series pre-whitening and local autocorrelation correction [20]. Co-registration to high-resolution and standard anatomical MR images was performed using FLIRT [18, 21]. BOLD percent signal changes were determined from the parameter estimates of the GLM and averaged across the activated region of each data set.

Higher-level (group) analyses were performed using FLAME (FMRIB's Local Analysis of Mixed Effects), which uses advanced Bayesian estimation techniques and applies Markov Chain Monte Carlo sampling (MCMC) for estimation of true random-effects variance and degrees of freedom [22]. Z (Gaussianised T/F) statistic images were thresholded using clusters determined by $Z > 3.5$ and a corrected cluster significance threshold of $P = 0.01$ [23-25]. Learning success, as reflected by retrieval performance, entered as a covariate in the statistical comparison between acquisition protocols to rule out variance due to differences in encoding efficacy. Residual intra-individual variance in the fMRI data was considered an effect of acquisition protocols. For statistical confirmation of the group results, we performed
paired t-tests and trend analyses in order to test for statistically significant differences between group statistical (T) maps as revealed from the different experimental conditions. Statistical maps from these paired t-tests were thresholded using clusters determined by $Z > 2.3$ and a corrected cluster significance threshold of $P = 0.05$. The statistical tests were restricted to the ROI under investigation. As different noise levels underlie the different acquisition strategies, we evaluated the statistical effect size ($Z$ values) rather than the physiological effect size (parameter estimates). The cross-subject analysis was therefore conducted on the t-statistic formed by dividing the first-level effect size by its standard deviation. We then performed an ordinary least-squares GLM analysis across subjects.

Comparison of acquisition protocols with regard to functional sensitivity was evaluated at group level since it composes the most relevant measure if an appropriate acquisition protocol needs to be selected for an fMRI study. Based on individual susceptibility distribution, different acquisition protocols, however, might provide better results for each subject individually. Moreover, large intra- and inter-subject variances, partially resulting from underlying learning, repetition and performance effects, prohibit a meaningful analysis on a single subject basis, as they may mask the pure effect from data acquisition strategy. The use of a random effects analysis and the inclusion of task performance as a confounding covariate hence yield more reliable information about the pure effect of data acquisition strategy.

RESULTS

Image quality

Figures 2-5 illustrate the effects of the different acquisition protocols on image quality with respect to susceptibility-related geometric image distortions. The SENSE acquisition yielded marked improvements in EPI image quality as apparent in reduced the left-right asymmetry, less bending and reduced stretching of the frontal sinuses along the anterior-posterior axis. Effective acquisition speed doubled and the sensitivity to geometric distortions dropped accordingly from 145.3 $\mu$m/Hz for conventional sshEPI to 79.0 $\mu$m/Hz for the twofold SENSE acquisition. The signal dropout typically found in orbitofrontal cortex was remarkably reduced at $TE = 22$ ms as compared to $TE = 35$ ms consistent with reduced signal dephasing in the susceptibility field gradients. At the same echo time, SENSE data yielded better signal
in this very region than conventional acquisition. Yet, benefits of higher SENSE factors were achieved mostly in the frontal lobes where the dephased signal was partly recovered. Otherwise only subtle further improvements were gained with higher acceleration factors $R = 2.4$, $2.7$, and $3.0$, and geometric distortions reduced to $79.0$, $64.3$, $56.8$, and $53.1 \, \mu m/Hz$ respectively. Representative data from all subjects are summarized in Figures 4 & 5, in order to illustrate the impact of and inter-individual differences in susceptibility-related artifacts depending on the individual anatomy and positioning within the magnetic field.

Figure 2: Sample EPI images taken from the functional time-series of one individual subject at five different slice locations. Consecutive rows show the conv-te22, conv-te35, sense-te22 and sense-te35 EPI images. Geometric distortions, apparent in bending and stretching of the images, were remarkably reduced with the SENSE acquisition (arrows). Grey ellipses outlined in the EPI images in the second row are to further support visual assessment and comparison of susceptibility-related image distortions. At lower echo time, signal dephasing and hence signal loss in field inhomogeneities was notably reduced and can be best appreciated in orbitofrontal cortex (rectangles). Moreover, at the same echo time, signal lost in the conventionally acquired images, can be partly recovered with the SENSE acquisition (circle).
Figure 3: Sample EPI images taken from the functional time-series of one individual subject at five different slice locations. Consecutive rows represent sense-2.0, sense-2.4, sense-2.7, and sense-3.0 data. Orbitofrontal cortex benefits from high SENSE acceleration factors (circled), since MR signal can be recovered in this typical drop-off region, which might otherwise be shifted outside the k-space matrix and hence is not sampled. Notably, geometric distortions were markedly reduced already at a SENSE acceleration factor of 2.0 and there was only minor further benefit from higher SENSE factors.
Figure 4: Sample slices from all subjects (columns) from the conv-te22, conv-te35, sense-te22, and sense-te35 functional time series (rows).

Figure 5: Sample slices from all subjects (columns) from the sense-2.0, sense-2.4, sense-2.7, and sense-3.0 functional time series (rows). The selected slices display transverse cuts through the medial temporal lobe.
Signal to Noise Ratio

SNR represents an important measure of fMRI signal stability. Figures 6a & b show SNR and noise maps for conv-te22, conv-te35, sense-te22 and sense-te35, and sense-2.0, sense-2.4, sense-2.7 and sense-3.0 data, respectively. Noise amplification was assessed by g-maps, which were computed from individual coil sensitivity maps. We furthermore add noise maps as derived from the standard deviation of the residuals from the GLM fit in order to visualize the spatial noise distribution in the functional data. SNR maps acquired with at TE = 35 ms showed notable SNR reduction as compared to the acquisition with TE = 22 ms. Moreover, conventional sshEPI data exhibited generally higher SNR than SENSE sshEPI data for the same echo time. However, signal gain in the frontal lobe was apparent in the SENSE sshEPI data, which was almost dephased in the conventional data (see also Figure 2). While SNR steadily decreased with higher SENSE acceleration, SNR heterogeneity increased and SNR dropped particularly in the MTL. Noise exhibited spatially dependent enhancement and areas of strong noise amplification increased with the SENSE acceleration. In contrast, signal was partly recovered in orbitofrontal cortex when measured with high SENSE acceleration, which might otherwise have been shifted outside the k-space acquisition matrix.

The theoretical SNR ratio between data acquired at different echo times, TE = 22 and 35 ms respectively, can be derived from T_2^* signal decay. The SNR ratio was approximated for gray matter (with T_2^* = 42 ms as derived from 1/T_2^* = 1/T_2 + k·B_0 and k = 4.8 (Ts)^{-1} and T_2 = 110 ms [26]) using Eq. [6] given in the appendix:

$$\frac{SNR_{TE=35ms}}{SNR_{TE=22ms}} = \frac{\exp(-35ms/T_2^*)}{\exp(-22ms/T_2^*)} \cdot \frac{T_{acq\text{TE}=35ms}}{T_{acq\text{TE}=22ms}}$$

The expected SNR ratio determined from theory between conv-te35 and conv-te22 is thus 81 % (Tacq_{conv-te35} = 52.8 ms, Tacq_{conv-te22} = 43.4 ms due to 82 % half Fourier acquisition), however was exceeded by the ratios determined in our data that revealed 93.1 ± 3.8 % in the MTL, 91.6 ± 2.9 % in the IFG and 88.6 ± 3.9 % in the FG. The commensurate SNR ratio for the sense-te35 and sense-te22 data (Tacq_{sense-te35} = Tacq_{sense-te22} = 28.7 ms), respectively, is 74 %. The average SNR ratio revealed in our data, however, was 82.7 ± 4.9 % in the MTL, 82.5 ± 8.3 % in the IFG and 77.9 ± 6.9 % in the FG.
Similarly, the SNR ratio between SENSE and conventional full Fourier encoded data can be estimated on the basis of reduced k-space sampling density and intrinsic noise amplification [4], as was derived in Eq. [6] in the appendix:

\[
\frac{SNR_{\text{SENSE}}}{SNR_{\text{FULL}}} = \frac{1}{g} \frac{\sqrt{T_{\text{acqSENSE}}}}{\sqrt{T_{\text{acqFULL}}}} \tag{2}
\]

which predicts a ratio of 74 % for an ideal geometry (\(g = 1\), \(T_{\text{acqSENSE-te35}} = 28.7\) ms, \(T_{\text{acqCONV-te35}} = 52.8\) ms). Taking into account the 82 % partial Fourier factor in the conv-te22 data (\(g = 1\), \(T_{\text{acqCONV-te22}} = 43.4\) ms, \(T_{\text{acqSENSE-te22}} = 28.7\) ms), the theoretical ratio between sense-te22 and conv-te22 is 81 %. The ratios determined from our data however were 86.8 ± 8.2 %, 88.2 ± 5.5 % and 97.8 ± 8.4 % (at TE = 22 ms) and 77.0 ± 6.4 %, 79.3 ± 8.6 % and 85.7 ± 7.0 % (at TE = 35 ms) in the MTL, IFG and FG respectively, therefore exceeding theoretical predictions. All SNR values are summarized in Table 1.

Predicted by Eq. [6] from the appendix, SNR ratios between SENSE acquisitions with different SENSE acceleration factors are 90 % (sense-24/sense-20), 85 % (sense-27/sense-20) and 82 % (sense-30/sense-20) under the assumption of an ideal geometry (\(g = 1\) and \(T_{\text{acqSENSE-20}} = 28.7\) ms, \(T_{\text{acqSENSE-24}} = 23.3\) ms, \(T_{\text{acqSENSE-27}} = 20.7\) ms, \(T_{\text{acqSENSE-30}} = 19.3\) ms). Ratios actually determined in our data, however, revealed 97.1 ± 8.4 %, 91.5 ± 10.5 % and 85.3 ± 9.3 % in the MTL, 96.3 ± 5.4 %, 87.9 ± 6.1 % and 81.0 ± 7.6 % in the IFG and 97.7 ± 8.0 %, 90.5 ± 8.4 % and 84.3 ± 7.4 % in the FG, respectively. Therefore all three regions of interest showed similar drops, however, ratios were beyond theoretical expectations. All SNR values are summarized in Table 2.
Figure 6: SNR (left) and noise maps (middle) for conv-te22, conv-te35, sense-te22, and sense-te35 images (a.) and for sense-2.0, sense-2.4, sense-2.7, and sense-3.0 images (b.). Shown on the right is the spatially varying noise amplification (g maps) in the SENSE sshEPI data as assessed from the coil sensitivity maps. Regions of intensive noise amplification increased with the SENSE acceleration factor - especially in the MTL at acceleration factor R = 3.0 and beyond. SNR and noise maps were computed from the motion-corrected, high-pass filtered and spatially unsmoothed data. All maps were scaled to the same intensity maximum to facilitate comparison across experimental conditions.
Table 1: SNR, size (# activated voxels) and mean significance (Z value) of clusters of activation, and mean BOLD signal change in non-te22, non-te35, sense-te22, and sense-te35 data, revealed in the MTL, IFG and FG. SNR was averaged across individual subjects (mean ± SD); while cluster size, mean BOLD signal change and mean significance were derived from the group results. MTL = medial temporal lobe, IFG = inferior frontal gyrus, FG = fusiform gyrus.

<table>
<thead>
<tr>
<th></th>
<th>SNR (unsmoothed)</th>
<th>SNR (smoothed)</th>
<th># act. voxel</th>
<th>BOLD [%]</th>
<th>Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>MTL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>non_22</td>
<td>84.7 ± 14.1</td>
<td>207.8 ± 38.1</td>
<td>1</td>
<td>0.57</td>
<td>3.54</td>
</tr>
<tr>
<td>non_35</td>
<td>78.5 ± 11.6</td>
<td>178.1 ± 27.9</td>
<td>19</td>
<td>0.53</td>
<td>3.63</td>
</tr>
<tr>
<td>sense_22</td>
<td>72.8 ± 9.6</td>
<td>178.2 ± 25.8</td>
<td>306</td>
<td>0.46</td>
<td>4.13</td>
</tr>
<tr>
<td>sense_35</td>
<td>60 ± 6.9</td>
<td>145.2 ± 18.2</td>
<td>324</td>
<td>0.46</td>
<td>3.98</td>
</tr>
<tr>
<td>IFG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>non_22</td>
<td>112.4 ± 8.4</td>
<td>249.7 ± 21.3</td>
<td>637</td>
<td>0.53</td>
<td>4.05</td>
</tr>
<tr>
<td>non_35</td>
<td>103.1 ± 9.7</td>
<td>214.1 ± 21.5</td>
<td>1903</td>
<td>0.56</td>
<td>4.26</td>
</tr>
<tr>
<td>sense_22</td>
<td>99.1 ± 9</td>
<td>224.5 ± 19.6</td>
<td>529</td>
<td>0.52</td>
<td>4.06</td>
</tr>
<tr>
<td>sense_35</td>
<td>81.4 ± 7.7</td>
<td>181.4 ± 16.9</td>
<td>1362</td>
<td>0.54</td>
<td>4.18</td>
</tr>
<tr>
<td>FG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>non_22</td>
<td>103.9 ± 8</td>
<td>223.8 ± 21.2</td>
<td>194</td>
<td>0.77</td>
<td>4.01</td>
</tr>
<tr>
<td>non_35</td>
<td>92.1 ± 8.8</td>
<td>188.7 ± 22.4</td>
<td>431</td>
<td>0.71</td>
<td>4.19</td>
</tr>
<tr>
<td>sense_22</td>
<td>101.4 ± 9.4</td>
<td>220.7 ± 21.3</td>
<td>767</td>
<td>0.74</td>
<td>4.22</td>
</tr>
<tr>
<td>sense_35</td>
<td>78.5 ± 4.3</td>
<td>175.1 ± 11</td>
<td>781</td>
<td>0.8</td>
<td>4.35</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>SNR (unsmoothed)</th>
<th>SNR (smoothed)</th>
<th># act. voxel</th>
<th>BOLD [%]</th>
<th>Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>MTL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sense_35/non_35</td>
<td>77 ± 6.4</td>
<td>82.2 ± 7.2</td>
<td>1705.3</td>
<td>86.8</td>
<td>109.6</td>
</tr>
<tr>
<td>sense_22/non_22</td>
<td>86.8 ± 8.2</td>
<td>86.8 ± 8.9</td>
<td>30600.0</td>
<td>80.7</td>
<td>116.7</td>
</tr>
<tr>
<td>sense_35/sense_22</td>
<td>82.7 ± 4.9</td>
<td>81.9 ± 5.8</td>
<td>105.9</td>
<td>100.0</td>
<td>96.4</td>
</tr>
<tr>
<td>non_35/non_22</td>
<td>93.1 ± 3.8</td>
<td>86.2 ± 4.4</td>
<td>1900.0</td>
<td>93.0</td>
<td>102.5</td>
</tr>
<tr>
<td>IFG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sense_35/non_35</td>
<td>79.3 ± 8.6</td>
<td>85.4 ± 10.2</td>
<td>71.6</td>
<td>96.4</td>
<td>98.1</td>
</tr>
<tr>
<td>sense_22/non_22</td>
<td>88.2 ± 5.5</td>
<td>90.1 ± 7.1</td>
<td>83.0</td>
<td>98.1</td>
<td>100.2</td>
</tr>
<tr>
<td>sense_35/sense_22</td>
<td>82.5 ± 8.3</td>
<td>81.2 ± 9.1</td>
<td>257.5</td>
<td>103.8</td>
<td>103.0</td>
</tr>
<tr>
<td>non_35/non_22</td>
<td>91.6 ± 2.9</td>
<td>85.7 ± 4.3</td>
<td>298.7</td>
<td>105.7</td>
<td>105.2</td>
</tr>
<tr>
<td>FG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sense_35/non_35</td>
<td>85.7 ± 7</td>
<td>93.8 ± 11.6</td>
<td>181.2</td>
<td>112.7</td>
<td>103.8</td>
</tr>
<tr>
<td>sense_22/non_22</td>
<td>97.8 ± 8.4</td>
<td>99.3 ± 12.1</td>
<td>395.4</td>
<td>96.1</td>
<td>105.2</td>
</tr>
<tr>
<td>sense_35/sense_22</td>
<td>77.9 ± 6.9</td>
<td>80 ± 9.4</td>
<td>101.8</td>
<td>108.1</td>
<td>103.1</td>
</tr>
<tr>
<td>non_35/non_22</td>
<td>88.6 ± 3.9</td>
<td>84.3 ± 5.9</td>
<td>222.2</td>
<td>92.2</td>
<td>104.5</td>
</tr>
</tbody>
</table>
Table 2: SNR, size (# activated voxels) and mean significance (Z value) of clusters of activation, and mean BOLD signal change in the sense-2.0, sense-2.4, sense-2.7 and sense-3.0 data revealed in the MTL, IFG and FG.

**Functional Activation Maps**

Our face-profession learning paradigm evoked significant activation in the left and right medial temporal lobe (MTL), the left and right inferior frontal gyri (IFG), the left and right fusiform (FG) and inferior occipital gyri (IOG) in the group analyses in all experimental conditions. Results from Experiment 1 group analysis revealed enhanced statistical power in the MTL, occipital lobe, and FG in the SENSE as compared to the conventional sshEPI data: We found more significant and larger clusters of activation bilaterally in anterior hippocampus and amygdala and in the left and right FG (Figure 7). Surprisingly, activation in the FG tended to be left lateralized in the conventional sshEPI data, while left and right FG were found activated in the SENSE sshEPI data. In contrast, clusters of activation in the IFG, and in particular Brodman area 45 and 47, were consistently larger in the conventional sshEPI
data. All results are summarized in Table 1. A paired t-test analysis performed in order to statistically confirm the differences observed in the group statistical maps, indeed confirmed our observations in the group data (Figure 8).

**Figure 7:** Group statistical maps overlaid on the mean, spatially normalized, anatomical MR image of all subjects in Experiment 1. Shown from the left to the right are coronal sections through the MTL, IFG and FG and from top to bottom conv-te22, conv-te35, sense-te22, and sense-te35 data, respectively. Activation within all regions of interest, except in the IFG, was better recovered with the SENSE compared to the conventional acquisition. At TE = 22 ms as compared to TE = 35 ms statistical power was reduced along with functional sensitivity as apparent in lower significance and smaller clusters of activation. Numbers indicate standard brain coordinates from the atlas of Talairach and Tournoix [27]. L = left, R = right.
Figure 8: Results from paired t-test analyses which were conducted to statistically confirm the observed differences between group statistical maps in Figure 7. Color coded are those regions that show significant differences (red = positive, blue = negative) in the statistical (T) maps as revealed from the different experimental conditions and were overlaid on the mean anatomy of all subjects. In all three regions of interest, MTL, IFG and FG more significant clusters of activation were revealed at TE = 35 ms as compared to TE = 22 ms and with the SENSE compared to the conventional acquisition (positive t-values), except for the right IFG that revealed better result with the conventional acquisition (negative t-values).

Intra-voxel dephasing progresses over time and can be diminished by shifting the acquisition towards lower echo times, which, however, is accompanied by comparable reduction in BOLD sensitivity (BS):

\[ I \propto \rho \cdot \exp\left( -\frac{TE}{T_2^*} \right) \]  \hspace{1cm} [3]

\[ BS \propto \frac{dl}{dT_2^*} = \frac{TE}{(T_2^*)^2} \cdot I \]  \hspace{1cm} [4]
Figure 9 shows a graphical representation of Eq. [4] for gray matter with field strength of 3 Tesla. Functional sensitivity maximum exhibits a reasonably flat plateau centered about the peak of the curve at \( TE = T_2^* \). Thus, a TE of 22 ms predicts a BOLD sensitivity reduction of only 16 % as compared to a TE of 35 ms.

Consistent with theoretical predictions, the comparison of group results at different echo times revealed reduced sensitivity to functional signal changes in the \( TE = 22 \) ms as compared to the \( TE = 35 \) ms data. Clusters of activation were significantly smaller at \( TE = 22 \) ms in all regions of interest for either image acquisition, with and without SENSE. The theoretically predicted difference in functional sensitivity was reflected in 2-3-fold larger clusters of activation in all three regions of interest in the conv-te35 than in the conv-te22 data. However, in the SENSE data they were larger only in the IFG (Table 1). In particular, we found significant activation only in the left IFG with \( TE = 22 \) ms, while the left and right IFG are found significantly activated at \( TE = 35 \) ms. We selected the common regions that showed significant activation at both \( TE = 22 \) ms and \( TE = 35 \) ms by using a mask created from a binary AND operation applied to significant clusters of activation in the statistical maps. Within these common regions, parameter estimates increased by 17.5 % in the FG and by 27.6 % in the IFG in the conv-te35 as compared to the conv-te22 group data. No spatial overlap was obtained for the MTL, since no significant activation was revealed in the MTL in the \( TE = 22 \) ms data. In the sense-te35 vs. the sense-te22 group results, parameter estimates increased by 6.0 % in the FG and by 20.2 % in the IFG, but did not change in the MTL.
The superior image quality in terms of reduced geometric artifact reduction, yet at comparable statistical power of the SENSE as compared to conventional sshEPI acquisition led us to further investigate the potential benefits and drawbacks of higher SENSE acceleration in a second fMRI study. Results from the group analysis of sense-2.0, sense-2.4, sense-2.7 and sense-3.0 data are summarized in Figures 10 & 11. Activation evoked in left and right hippocampus and amygdala yielded the highest significance and largest clusters of activation in the sense-2.0 data. In the MTL activation showed a decreasing trend in terms of cluster size and statistical power with intermediate SENSE acceleration, while a strong functional response was again revealed in the sense-3.0 data. Similarly, cluster size and significance within the left IFG were largest in the sense-2.0 data and show a decreasing trend with increasing SENSE acceleration. Moreover, while the IFG was activated bilaterally in the sense-2.7 and sense-3.0 data, activation was lateralized to the left side in the sense-2.0 and sense-2.4 data. In contrast, activation along the left and right FG and IOG showed increasing cluster size and significance with increasing SENSE acceleration factors. These findings suggest that, surprisingly, neocortical areas (not IFG) benefited from high SENSE factors, while activation in the MTL and IFG were better assessed with modest SENSE acceleration. Results from the group analysis are summarized in Table 2. Mean Z values were not statistically different between SENSE acquisitions in neither region of interest, while the size of clusters of activation showed significant trend between the measurement conditions. These results were confirmed by paired t-tests and trend analyses as shown in Figure 11.
Figure 10: Group statistical maps overlaid on the mean, spatially normalized, anatomical MR image of all subjects in Experiment 2. Shown from the left to the right are coronal sections through the MTL, IFG and FG and from top to bottom sense-2.0, sense-2.4, sense-2.7, and sense-3.0 data, respectively. Activation in the MTL was equally reproduced in the sense-20 and sense-30 data, however was markedly reduced in the sense-24 and sense-27 data. Activation within the left IFG was best revealed with sense-2.0, while activation in posterior neocortex, along the FG and IOG was surprisingly better recovered at high SENSE factors. Numbers indicate standard brain coordinates from the atlas of Talairach and Tournoux [27]. L = left, R = right.
I'iguie 11: Results from paired t-test and trend analyses which were conducted to statistically confirm the observed differences between group statistical maps shown in Figure 10. Color coded are those regions that show significant differences (red = positive, blue = negative) in the statistical (T) maps as revealed from the different experimental conditions and were overlaid on the mean anatomy of all subjects. Consistent with the results observed in the group analyses, these data show consistently higher significance within the MTL in the acquisition in the sense-20 as compared to the sense-24 and sense-27 data, while it is comparable between sense-20 and sense-30 data (better in the superior, worse in the inferior section of the left hippocampus). This 'U-shaped' trend was significant in the left hippocampus. The IFG, however, shows declining significance with increasing SENSE factors (negative in the linear increasing trend analysis), while significance of the activation within the FG increases with increasing SENSE acceleration (positive response in the linear-increase trend analysis, negative response in the sense-20>sense-24, sense-20>sense-27 and sense-20>sense-30 paired t-tests).

Figure 11: Results from paired t-test and trend analyses which were conducted to statistically confirm the observed differences between group statistical maps shown in Figure 10. Color coded are those regions that show significant differences (red = positive, blue = negative) in the statistical (T) maps as revealed from the different experimental conditions and were overlaid on the mean anatomy of all subjects. Consistent with the results observed in the group analyses, these data show consistently higher significance within the MTL in the acquisition in the sense-20 as compared to the sense-24 and sense-27 data, while it is comparable between sense-20 and sense-30 data (better in the superior, worse in the inferior section of the left hippocampus). This 'U-shaped' trend was significant in the left hippocampus. The IFG, however, shows declining significance with increasing SENSE factors (negative in the linear increasing trend analysis), while significance of the activation within the FG increases with increasing SENSE acceleration (positive response in the linear-increase trend analysis, negative response in the sense-20>sense-24, sense-20>sense-27 and sense-20>sense-30 paired t-tests).
DISCUSSION

We performed two fMRI studies in order to identify suitable imaging parameters for fMRI in the medial temporal lobe at 3 Tesla field strength. A paired-associate learning paradigm was used in both studies. In Experiment 1, SENSE acquisition was compared to conventional image acquisition at both TE = 22 and 35 ms. In Experiment 2, we compared between functional data acquired with SENSE acceleration factors 2.0, 2.4, 2.7, and 3.0 respectively.

Our data are in accordance with the theoretically predicted benefits from parallel imaging [28], as became apparent in reduced geometric image distortions when compared to conventional sshEPI data. Although not quantified in this study, reduction of susceptibility artifacts in the functional images might further contribute to more reliable anatomical assignment of resulting clusters of activation and transformation into standard brain space. It should be noted that the same benefits underlie segmented acquisition strategies [29], which, however, require multiple shots for each complete image and hence are more time consuming which might not be affordable when functional time series are to be obtained with high temporal precision and when whole brain coverage is mandatory. Furthermore, they might be sensitive to shot-to-shot inconsistencies such as provoked by subject motion. It hence becomes clearly apparent that parallel imaging provides a far more elegant solution, since spatial and temporal resolution are preserved.

Reduction of geometric image distortions was most apparent between conventional and SENSE sshEPI at R = 2.0. Corresponding to the reduction of acquisition time, sensitivity to susceptibility-related distortions was reduced by a factor of two, while only a moderate (1.5 fold) further reduction was achieved in the sense-30 as compared to the sense-20 data, resulting in only subtle further reduction of geometric image distortions.

In Experiment 1 SNR in the TE = 35 ms as compared to TE = 22 ms data was larger than straight-forward theory predicts (Eq. [1]) for both, conventional and SENSE data. However, our estimation of SNR was based on gray matter only, and a CSF contribution, however, might shift the ratio up to higher values. Comparison between conventional and SENSE acquisition, as well as comparison between different SENSE acceleration factors in Experiment 2 revealed that even with the assumption of an ideal geometry (g = 1), the SNR was less reduced as predicted from undersampling alone. Importantly, Eq. [1] holds strictly only for intrinsic ('image') SNR while we took also temporal signal fluctuations into account.
Since physiological noise emerges as the dominant noise source in BOLD-fMRI at high field strength readily limiting maximum attainable SNR \[30\], it is conceivable that with an estimation based on pure 'image SNR', SNR drawback is likely to be overestimated \[5\]. Moreover, Eq. \([6]\) does not take into account that signal might be recovered in typical dropout regions such as in orbitofrontal and medial temporal cortex with SENSE as compared to conventional image acquisition, which might also partly account for the larger SNR ratios in our the experimental data. Likewise, different T2*-decay during the acquisition train is not included and it should furthermore be noted that with the shorter echo train length the distribution of effective echo times along phase-encoding might be centered more closely about the nominal echo time at \(k = 0\) and hence BOLD-weighting be more uniform in k-space.

Although we analyzed the functional data within the whole brain, we focused especially on activation in three regions of interest - the MTL, IFG and FG, where, consistent with previous studies, our learning paradigm evoked significant activation \[15\]. While the hippocampus is essential for associative memory processing \[31\] and has correspondingly been shown activated during encoding of novel faces and face-name associations \[32, 33\], visual perception of faces is mediated by a distributed cortical network \[34, 35\] including the IOG and FG where invariant aspects of faces such as configural information and the unique identity are processed \[34, 36\]. The IFG mediates the perception of emotional facial expressions and is involved in visual working memory and imagery \[32, 37\].

The MTL was of particular interest in the present study since reconstruction-related noise enhancement is most critical in this very region and, furthermore, evoked functional response is rather small. This raised the question, as to whether at all fMRI in the MTL can benefit from parallel imaging.

In Experiment 1 a TE = 22 ms was accompanied by a reduction of susceptibility related signal dephasing especially in orbitofrontal cortex, reduced functional sensitivity concomitantly led to a degradation in statistical power, which was apparent in a reduction of the size of significant clusters activation and reduced BOLD signal change in the common regions that showed significant activation at both echo times. Statistical significance decreased correspondingly, although to smaller degree, which might be attributable to lower SNR at TE = 35 ms. The reduction of functional sensitivity was best seen in the IFG and FG but was less apparent in the MTL, where physiological effect size and statistical power did not changed
considerably between TE = 22 ms and TE = 35 ms. Previous studies have shown that the effective echo time is affected by susceptibility gradients, and hence local BOLD sensitivity might differ between different regions of the brain such as the MTL and IFG and FG in our study [3]. Most remarkably, however, our SENSE data revealed more significant and larger clusters of activation in all regions of interest, except in the IFG, as compared to the conventional data.

In Experiment 2 comparison between different SENSE acceleration factors revealed that activation in the MTL was reliably detected in the sense-20 data, but did not reveal comparable statistical significance and cluster size at intermediate SENSE acceleration factors, yet emerged again in the sense-30 data. Notably, the poor statistical outcome in the sense-24 and sense-27 data was also observed in a pilot study conducted with 8 subjects, and is consistent with lower intrinsic SNR as compared to the sense-20 data. Based on the present study, however, we are not able to provide strong and conclusive arguments for the kind of 'U-shaped' trend that revealed more significant and larger clusters of activation at both extreme (R = 2.0 and R = 3.0) as compared to the intermediate SENSE acceleration factors (R = 2.4 and R = 2.7). Yet, the theoretical framework, predicts a decreasing trend, due to progressive undersampling and noise amplification associated with the data reconstruction, especially in the MTL. Therefore, the technical or even physiological underpinnings for these adverse observations necessitate further investigation. Still, we are going to present one possible explanation further down.

Activation in the left IFG concomitantly decreased with increasing SENSE acceleration and we did not observe comparable recovery of functional activation in den sense-30 data, as we did in the MTL. In contrast, and surprisingly, activation in the FG and IOG however exhibited more significant and larger clusters of activation in the high as compared to the low SENSE acceleration data. This effect was also present in the visual stimulation data shown in Figure 12 and described below. Unlike the MTL, occipital cortex, however, is not as much affected by intrinsic noise amplification (g maps in Figure 6). This inverse trend found in the FG and IOG might possibly originate from shorter acquisition duration associated with the SENSE acquisition, but certainly requires further investigation. While previous studies have shown that significance of activation (t-values) in primary sensory regions does not as much decline as would be expected based on reduced image SNR alone, but was rather similar to the loss
predicted by a model that took physiologic noise into account [5], and actual increase of t-values was not yet reported.

Arguably and although not consistently observed, residual foldover from the reconstruction process might have caused functional signal to be warped from the occipital cortex into the MTL. Figure 12a represents mean anatomical images, which were 2.0-, 2.4-, 2.7- and 3.0-fold undersampled corresponding to the SENSE acceleration. Apparently, the foldover effect becomes increasingly critical in the MTL at and beyond threefold undersampling. Supposedly, it might occur that even if a residual aliasing artifact is not visible in the images themselves, a small, warped functional signal might yet be sufficient to induce a statistical effect. To investigate this effect, we performed a basic visual fMRI experiment in four volunteers, using a task that reliably evokes a functional response in occipital cortex but does not recruit the MTL. Simulation consisted of an 8 Hz reversing black/white checkerboard stimulus that alternated with a fixation cross in four repetitive 20 s cycles. We used the same SENSE acquisition protocols and aligned the slices along the calcarine tissue. The functional activation maps from this experiment are represented in Fig. 12b, showing that with increasing SENSE acceleration more functional activation was revealed in the MTL, which was not present in the sense-20 and sense-24 data. These data suggest that some functional signal might indeed have been warped from the occipital cortex into the MTL.

Finally we want to emphasize, that our findings represent a general trend specific for functional imaging in the medial temporal lobe, the inferior frontal and fusiform gyri for groups of subjects. However, the optimum acquisition protocol might be different for each individual subject and in other ROIs depending on the individual brain anatomy and positioning within the MR scanner [38].
Figure 12: a. Foldover effects resulting from undersampling in k-space with the SENSE acquisition. Mean anatomical images were undersampled by factors of 2.0, 2.4, 2.7, and 3.0 corresponding to the different SENSE acceleration factors, and demonstrate that these become increasingly critical in the MTL at and beyond threefold undersampling. b. Functional activation maps obtained from a basic visual fMRI experiment raise the possibility that residual aliasing might have caused functional signal to be warped from the occipital cortex into the MTL. Notably, with increasing SENSE acceleration more functional activation is observed in the MTL (rectangle), which is not present in the data acquired with low SENSE acceleration factors $R = 2.0$ and 2.4.
CONCLUSION

In this work we presented results from various imaging protocols that adopt Sensitivity Encoded (SENSE) parallel imaging for fMRI in the medial temporal activation at 3 Tesla. Susceptibility-related image distortions were markedly reduced in the SENSE as compared to the conventionally full Fourier encoded sshEPI images. While data acquired with SENSE acceleration factors beyond 2 revealed only subtle further reduction of geometric distortions, enhanced encoding speed was at remarkable expense of statistical significance in the functional activation maps in the medial temporal lobe except for an unexpected improvement at extreme acceleration factor (R = 3.0). Our data suggest that sshEPI acquisition with a moderate SENSE acceleration factor of R = 2.0 is suitable for fMRI experiments aiming at the detection of medial temporal activation.

Acknowledgments

We thank Dr. J. Tsao and Prof. A. Ishai for valuable discussions and critical reading of the manuscript and Prof. S.M. Smith from FMRIB for his instant support during data analysis. The authors also wish to thank for the continuing support of Philips Medical systems and the financial support by ETH Zurich (Strategic Excellence Project TH 7/02-02). K. Henke was supported by a grant from the Neuroscience Center Zurich and a grant from the Swiss National Science Foundation (3100-067114.01).

APPENDIX

The induction signal detected in the receiver coils is proportional to the voxel volume $V$, the number of samples, i.e. the product of the number of phase-encoding steps $N_p$ and the number of frequency encoding steps $N_f$, whereas the noise scales with the square root of the number of samples and with the acquisition bandwidth of the receiver. Thus, for the SNR follows:

$$\text{SNR} \propto \frac{V \cdot N_p \cdot N_f}{\sqrt{bw \cdot N_p \cdot N_f}}$$

[5]
and the acquisition bandwidth can be approximated from:

\[ bw \sim \frac{N_F \cdot N_F}{T_{acq}} \]

with \( T_{acq} \) representing the total acquisition duration. The application of Sensitivity Encoding introduces spatially varying noise amplification, described by the so called geometry factor \( g \) [4] that emerges from non-unitary operations in image reconstruction, and typically results in increased noise in the center of the reconstructed images.

If data acquisition is performed at different echo times, an additional term for \( T_2^* \) decay needs to be included in the SNR calculation:

\[ SNR \approx \frac{V \cdot \sqrt{T_{acq}}}{g} \cdot \exp\left(-\frac{TE}{T_2^*}\right) \quad [6] \]

REFERENCES:


17. www.fmrib.ox.ac.uk/fsl.


ABSTRACT

In this study we compared fMRI activation measured with gradient- and spin-echo based fMRI during visual perception of faces, which is mediated by neural activation within a distributed cortical network. With both fMRI techniques bilateral activation was observed in multiple regions including the inferior occipital gyrus, fusiform gyrus, superior temporal sulcus, amygdala, inferior frontal gyrus, and orbitofrontal cortex. When compared with the gradient-echo sequence, activation measured with the spin-echo sequence was significantly reduced. This decrease was manifested by smaller cluster size, lower statistical significance, smaller amplitude of the fMRI signal, and smaller number of subjects who showed activation in all face-responsive regions. In orbitofrontal cortex, a region prone to susceptibility-related signal dephasing, the spin-echo acquisition considerably restored the signal, but did not reveal stronger activation when compared with the gradient-echo acquisition. Our data indicate that optimized GE sequences that reduce susceptibility artefacts are sufficient to detect activation in regions such as the orbitofrontal cortex.
INTRODUCTION

Functional MRI based on blood oxygenation level-dependent (BOLD) contrast has become the most commonly used methodology for mapping human brain function. To localize foci of activation with high spatial resolution, it is essential to minimize the macrovascular contribution of large draining veins that may be distant from the actual site of neuronal activity. The vast majority of cognitive studies have employed gradient-echo (GE) sequences, which are sensitive to both large vessels and small capillaries. Spin-echo (SE) sequences, in contrast, eliminate non-specific contributions from extravascular BOLD signal changes in the vicinity of large venous vessels, and are less sensitive to susceptibility-related signal dropout. It has been suggested that with SE contrast, the functional response can be mapped with higher spatial specificity [1-3], provided that intravascular signal changes are eliminated, and that activation in regions affected by strong susceptibility gradients, such as the orbitofrontal cortex, can be localized [4].

Previous studies have shown that SE-based fMRI is a useful tool to map functional activation in primary visual and motor cortices [5, 6, 1, 7-11, 2, 12, 3]. Moreover, it has been shown that SE contrast at 3 Tesla can be applied to cognitive fMRI studies [4]. In these studies, larger clusters of activation and higher functional signal changes were reported in the GE- as compared to the SE-EPI data. These differences were attributed to greater microvascular selectivity of the extra-vascular BOLD signal in SE-based fMRI, as non-specific extravascular signal changes in the vicinity of large vessels were rephased.

The functional contrast obtained with SE-based fMRI differs from the one underlying GE-based fMRI. In an SE experiment, static dephasing effects that contribute to the BOLD signal are refocused. This results in reduced functional signal by a factor of 2 – 4, depending on the underlying vascular architecture [13]. Although providing lower functional contrast, SE-based fMRI can be beneficial in regions where signal is largely dephased in a GE acquisition (i.e., when functional signal cannot be detected). We hypothesized that in regions that suffer from susceptibility-related dropout, such as orbitofrontal cortex, the MRI signal would be restored with an SE sequence and hence a functional signal detected.

In this study we directly compared activation measured with GE- and SE-EPI sequences during visual perception. We chose a simple task, namely attentive viewing, and various stimulus formats (line drawings of unfamiliar faces and photographs of unfamiliar, famous,
and emotional faces), in order to localize activation within multiple, bilateral face-responsive regions in extrastriate cortex, the limbic system, and prefrontal cortex [14-16]. For each subject we identified a set of cortical regions activated during face perception, as measured with both GE- and SE-EPI sequences, in the same experimental session and with the same statistical threshold.

We found bilateral activation within a network of face-responsive regions, including the inferior occipital gyrus (IOG), fusiform gyrus (FG), superior temporal sulcus (STS), amygdala, inferior frontal gyrus (IFG), and orbitofrontal cortex (OFC). Activation measured with both GE- and SE-EPI sequences was found in the same face-responsive regions, however the spatial extent of the activation, statistical significance, and amplitude of the fMRI signal were significantly smaller in the SE-EPI data.

**EXPERIMENTAL PROCEDURES**

*Subjects:* Thirteen healthy volunteers (8 males, 5 females, mean age 26 ± 4 yr) with normal vision participated in the study. All subjects gave written informed consent for the procedure.

*Stimuli and Paradigm:* Subjects were presented with four different types of face stimuli: black and white line drawings of unfamiliar faces and gray scale photographs of unfamiliar, famous, and emotional faces. Phase scrambled versions of these faces were used as visual baseline. The scrambled pictures were generated by randomizing the phase information after Fourier transformation using an in-house Matlab script. Each stimulus was presented for 3 sec. Each time series included three alternating epochs of faces (36 sec) and scrambled faces (24 sec). Five runs (line drawings, famous, emotional, and 2 runs with unfamiliar faces) were collected for each subject. The order of stimulus formats was randomized and counterbalanced. Stimuli were generated using SCOPE V2.5.4 (Max R. Duersteler, University Hospital Zurich, Switzerland) and were projected with a magnetically shielded LCD video projector onto a translucent screen. The subject viewed the screen by a mirror system.

*Data Acquisition:* Data were collected using a 3T Philips Intera whole body MR scanner (Philips Medical Systems, Best, The Netherlands) equipped with a transmit-receive body coil and a commercial eight-element head receiver array (MRI Devices Corporation, Waukesha
Functional activation was measured using a GE-EPI sequence (TE = 35 ms, 39 transverse slices) and an SE-EPI sequence (TE = 75 ms, 27 transverse slices). Slice geometry and orientation were otherwise the same for both acquisitions. A spatial resolution of 2.3 x 2.3 x 3 mm³ (acquisition matrix 96 x 96) was obtained using Sensitivity Encoding (SENSE) with an acceleration factor of 2.0. Parallel imaging techniques, such as SENSE, enable faster encoding by using different spatial coil sensitivity profiles for image reconstruction [17]. Thus, for a given spatial resolution, susceptibility related artefacts such as image distortion and blurring can be reduced by shortening the required echo train length [18-22]. The acquisition time for each slice was equal in both sequences, T_{acq} = 38 ms. With typical values for T₂* and T₂, which are in the order of the acquisition time or larger, line broadening due to T₂*- and T₂-decay is in the range of or smaller than the full-width-at-half-maximum (FWHM) of the sampling point spread function. Other functional imaging parameters were FOV = 220 mm, TR = 3000 ms and θ = 82°. High-resolution spoiled gradient recalled echo structural images were obtained with 1 x 1 x 0.8 mm³ spatial resolution, acquisition matrix 224 x 224, TE = 2.30 ms, TR = 20 ms, θ = 20°. These T₁-weighted images provided detailed anatomical information for the region of interest (ROI) analysis.

Data Analysis: Data were analyzed using the SPM2 software (www.fil.ion.ucl.ac.uk/spm/). All volumes were realigned to the first volume, corrected for motion artefacts, mean-adjusted by proportional scaling, normalized into standard stereotactic space (template provided by the Montreal Neurological Institute), and spatially smoothed using a 5 mm FWHM Gaussian kernel. The time series were high-pass filtered to eliminate low-frequency components (filter width 128 sec) and adjusted for systematic differences across trials. The main effect of faces (i.e., the response to faces versus the response to scrambled faces) was analyzed using a linear convolution model with an assumed hemodynamic response function [23, 24]. Clusters were selected that showed a significant effect (p < 0.001, uncorrected, with cluster size of 4 or more contiguous voxels). Statistical parametric maps indicating the main effect of faces were used to identify a set of ROIs for each subject, including bilaterally the inferior occipital gyrus (IOG), fusiform gyrus (FG), superior temporal sulcus (STS), amygdala, and the inferior frontal gyrus (IFG). Activation in the orbitofrontal cortex (OFC) was found medially and the clusters were small, the ROI analysis for this region therefore included both hemispheres. The anatomical locations of these clusters were determined by superimposing the statistical maps on the coplanar high-resolution structural images. ROI masks were created using MRICro (www.psychology.nottingham.ac.uk). For each subject and each ROI, a mean time series,
averaged across activated voxels in a region and across all repetitions for each stimulus format, was calculated. These means were used for between-subjects random-effects analyses. Separate repeated measures ANOVAs analyzed the effect of contrast (GE- and SE-EPI) in each region and each hemisphere. We also performed a conjunction analysis on the common regions activated with both GE and SE sequences (the IOG, FG, and STS).

RESULTS

Image Quality and Susceptibility Effects

Figure 1 illustrates GE- and SE-EPI images taken from four representative subjects. While severe signal voids in orbitofrontal cortex and around the inner ear canal were evident in the GE images, the MR signal in these susceptible regions was refocused with the SE-EPI sequence. Additionally, signal dropout in the vicinity of large venous blood vessels observed with the GE sequence, was recovered with the SE-EPI (e.g., subject 4).

![Figure 1: Image quality and susceptibility related distortions in gradient- and spin-echo EPI images. Transverse slices taken from four individuals show representative GE- and SE-EPI functional images. While GE-EPI images showed severe signal dropout in orbitofrontal cortex and around the inner ear canals, dephased signal was refocused in the SE-EPI images.](image)

Activation Within a Network of Face-Responsive Regions

With both GE- and SE-based fMRI, we found activation in response to faces within multiple, bilateral regions (Figure 2). Attentive viewing of faces, as compared with scrambled faces,
evoked significant responses in the IOG, FG, STS, amygdala, IFG, and OFC. We found bilateral activation within these regions (see Table 1 for cluster size and brain atlas coordinates [25]), however stronger responses were observed in the right than in the left hemisphere. In all face-responsive regions, the spatial extent of the activation was larger in the right hemisphere (GE-EPI p < 0.001, SE-EPI p < 0.01). Additionally, within the IOG and FG, higher t-values were found in the right hemisphere in the GE-EPI data (p < 0.001).

Figure 2: A network of face-responsive regions. Shown from the left to right are coronal sections taken from two individuals, illustrating activation measured with GE- and SE-EPI in the inferior occipital gyrus (IOG), fusiform gyrus (FG), superior temporal sulcus (STS), amygdala (AMG), inferior frontal gyrus (IFG), and orbitofrontal cortex (OFC). Statistical maps indicate the main effect of faces (p < 0.001, clusters of 4 or more contiguous voxels).
### GE-EPI

<table>
<thead>
<tr>
<th>Region</th>
<th>N</th>
<th>mean cluster size</th>
<th>mean coordinates</th>
<th>mean t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>[cm³]</td>
<td>x</td>
<td>y</td>
</tr>
<tr>
<td>L. IOG</td>
<td>13</td>
<td>3.05 (0.62)</td>
<td>-41 (1)</td>
<td>-81 (1)</td>
</tr>
<tr>
<td>R. IOG</td>
<td>13</td>
<td>3.98 (0.67)</td>
<td>41 (2)</td>
<td>-82 (2)</td>
</tr>
<tr>
<td>L. FG</td>
<td>13</td>
<td>2.92 (0.41)</td>
<td>-42 (1)</td>
<td>-58 (2)</td>
</tr>
<tr>
<td>R. FG</td>
<td>13</td>
<td>4.42 (0.55)</td>
<td>40 (1)</td>
<td>-57 (2)</td>
</tr>
<tr>
<td>L. STS</td>
<td>10</td>
<td>0.59 (0.19)</td>
<td>-54 (3)</td>
<td>-48 (4)</td>
</tr>
<tr>
<td>R. STS</td>
<td>13</td>
<td>0.80 (0.22)</td>
<td>53 (1)</td>
<td>-45 (3)</td>
</tr>
<tr>
<td>L. AMG</td>
<td>8</td>
<td>0.64 (0.17)</td>
<td>-18 (2)</td>
<td>-7 (1)</td>
</tr>
<tr>
<td>R. AMG</td>
<td>11</td>
<td>0.65 (0.16)</td>
<td>19 (1)</td>
<td>-7 (1)</td>
</tr>
<tr>
<td>L. IFG</td>
<td>7</td>
<td>0.87 (0.29)</td>
<td>-47 (2)</td>
<td>19 (2)</td>
</tr>
<tr>
<td>R. IFG</td>
<td>11</td>
<td>1.41 (0.46)</td>
<td>52 (1)</td>
<td>23 (3)</td>
</tr>
<tr>
<td>B. OFC</td>
<td>8</td>
<td>0.79 (0.34)</td>
<td>-2 (2)</td>
<td>50 (1)</td>
</tr>
</tbody>
</table>

### SE-EPI

<table>
<thead>
<tr>
<th>Region</th>
<th>N</th>
<th>mean cluster size</th>
<th>mean coordinates</th>
<th>mean t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>[cm³]</td>
<td>x</td>
<td>y</td>
</tr>
<tr>
<td>L. IOG</td>
<td>13</td>
<td>1.43 (0.44)</td>
<td>-43 (1)</td>
<td>-78 (1)</td>
</tr>
<tr>
<td>R. IOG</td>
<td>10</td>
<td>1.87 (0.78)</td>
<td>39 (1)</td>
<td>-81 (2)</td>
</tr>
<tr>
<td>L. FG</td>
<td>12</td>
<td>1.13 (0.21)</td>
<td>-39 (2)</td>
<td>-55 (3)</td>
</tr>
<tr>
<td>R. FG</td>
<td>12</td>
<td>1.90 (0.41)</td>
<td>38 (1)</td>
<td>-56 (2)</td>
</tr>
<tr>
<td>L. STS</td>
<td>4</td>
<td>0.16 (0.07)</td>
<td>-56 (5)</td>
<td>-40 (9)</td>
</tr>
<tr>
<td>R. STS</td>
<td>6</td>
<td>0.23 (0.10)</td>
<td>54 (1)</td>
<td>-39 (3)</td>
</tr>
<tr>
<td>L. AMG</td>
<td>5</td>
<td>0.14 (0.03)</td>
<td>-19 (2)</td>
<td>-3 (3)</td>
</tr>
<tr>
<td>R. AMG</td>
<td>4</td>
<td>0.17 (0.07)</td>
<td>15 (2)</td>
<td>-6 (1)</td>
</tr>
<tr>
<td>L. IFG</td>
<td>4</td>
<td>0.13 (0.07)</td>
<td>-46 (2)</td>
<td>25 (3)</td>
</tr>
<tr>
<td>R. IFG</td>
<td>6</td>
<td>0.67 (0.29)</td>
<td>52 (1)</td>
<td>24 (3)</td>
</tr>
<tr>
<td>B. OFC</td>
<td>9</td>
<td>0.85 (0.43)</td>
<td>-2 (1)</td>
<td>42 (2)</td>
</tr>
</tbody>
</table>

**Table 1:** Main effect of faces as measured with gradient- and spin-echo EPI. N indicates the number of subjects who showed significant activation (p < 0.001, clusters of 4 or more contiguous voxels). Volumes and t-values were calculated before spatial normalization. Coordinates are in the normalized space of the Talairach and Tournoux brain atlas. For each region, mean volume, mean t-value and mean coordinates were averaged across all subjects. SEM are indicated in parentheses. L = left, R = right, M = medial.

**Gradient- versus Spin-Echo Based fMRI**

For most subjects, face-responsive regions detected with the GE-EPI sequence were also revealed with the SE-EPI sequence. The SE-EPI statistical maps, however, revealed activation in smaller clusters as compared with the activation measured with the GE-EPI sequence with the same statistical threshold (Figure 2). Both cluster size (Figure 3) and t-values (Figure 4) were significantly smaller in the SE-EPI data as compared with the GE-EPI data (p < 0.0001 and p < 0.0001, respectively). The larger differences in cluster size and t-value between the GE- and SE-EPI sequences were observed in the visual cortex, namely in the IOG and FG.
Differential Responses to Various Face Formats

To compare the differential activation evoked by the various face stimuli as measured with the GE- and SE-EPI sequences (Table 2), we performed a conjunction analysis. In most subjects, the common regions showing activation with both techniques were the IOG (13 subjects showed activation in the left hemisphere, 12 in the right), FG (12 subjects showed bilateral activation), and the STS (5 subjects in the left hemisphere, 8 in the right). Within these common regions, we found smaller amplitudes of the fMRI signal in the SE-EPI data (Figure 5, p < 0.001). Regardless of the fMRI sequence, famous and emotional faces evoked stronger responses than unfamiliar faces bilaterally in the IOG and FG, and in the right STS (GE-EPI p < 0.001, SE-EPI p < 0.01).
### GE-EPI

<table>
<thead>
<tr>
<th>Region</th>
<th>Line drawings</th>
<th>Unfamiliar faces</th>
<th>Famous faces</th>
<th>Emotional faces</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>mean cluster size [cm$^3$]</td>
<td>mean t-value</td>
<td>N</td>
</tr>
<tr>
<td>L. IOG</td>
<td>11</td>
<td>1.75 (0.42)</td>
<td>4.13 (0.17)</td>
<td>12</td>
</tr>
<tr>
<td>R. IOG</td>
<td>12</td>
<td>1.82 (0.33)</td>
<td>4.30 (0.22)</td>
<td>11</td>
</tr>
<tr>
<td>L. FG</td>
<td>13</td>
<td>1.15 (0.33)</td>
<td>4.02 (0.15)</td>
<td>12</td>
</tr>
<tr>
<td>R. FG</td>
<td>13</td>
<td>2.11 (0.45)</td>
<td>4.18 (0.18)</td>
<td>13</td>
</tr>
<tr>
<td>L. STS</td>
<td>4</td>
<td>0.10 (0.03)</td>
<td>3.45 (0.12)</td>
<td>6</td>
</tr>
<tr>
<td>R. STS</td>
<td>5</td>
<td>0.14 (0.09)</td>
<td>3.52 (0.07)</td>
<td>7</td>
</tr>
<tr>
<td>L. AMG</td>
<td>2</td>
<td>0.19 (0.10)</td>
<td>3.53 (0.04)</td>
<td>4</td>
</tr>
<tr>
<td>R. AMG</td>
<td>3</td>
<td>0.12 (0.06)</td>
<td>3.46 (0.11)</td>
<td>5</td>
</tr>
<tr>
<td>L. IFG</td>
<td>6</td>
<td>0.82 (0.64)</td>
<td>3.52 (0.12)</td>
<td>4</td>
</tr>
<tr>
<td>R. IFG</td>
<td>8</td>
<td>1.11 (0.59)</td>
<td>3.68 (0.15)</td>
<td>6</td>
</tr>
<tr>
<td>B. OFC</td>
<td>5</td>
<td>0.40 (0.24)</td>
<td>3.67 (0.08)</td>
<td>4</td>
</tr>
</tbody>
</table>

### SE-EPI

<table>
<thead>
<tr>
<th>Region</th>
<th>Line drawings</th>
<th>Unfamiliar faces</th>
<th>Famous faces</th>
<th>Emotional faces</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>mean cluster size [cm$^3$]</td>
<td>mean t-value</td>
<td>N</td>
</tr>
<tr>
<td>L. IOG</td>
<td>7</td>
<td>0.58 (0.19)</td>
<td>3.67 (0.10)</td>
<td>9</td>
</tr>
<tr>
<td>R. IOG</td>
<td>7</td>
<td>0.61 (0.30)</td>
<td>3.65 (0.11)</td>
<td>9</td>
</tr>
<tr>
<td>L. FG</td>
<td>7</td>
<td>0.34 (0.12)</td>
<td>3.71 (0.16)</td>
<td>9</td>
</tr>
<tr>
<td>R. FG</td>
<td>8</td>
<td>0.65 (0.26)</td>
<td>3.68 (0.07)</td>
<td>11</td>
</tr>
<tr>
<td>L. STS</td>
<td>0</td>
<td>---- ----</td>
<td>---- ----</td>
<td>5</td>
</tr>
<tr>
<td>R. STS</td>
<td>1</td>
<td>0.04 (0.00)</td>
<td>3.23 (0.00)</td>
<td>3</td>
</tr>
<tr>
<td>L. AMG</td>
<td>1</td>
<td>0.04 (0.00)</td>
<td>3.36 (0.00)</td>
<td>0</td>
</tr>
<tr>
<td>R. AMG</td>
<td>1</td>
<td>0.10 (0.00)</td>
<td>3.46 (0.00)</td>
<td>3</td>
</tr>
<tr>
<td>L. IFG</td>
<td>1</td>
<td>0.07 (0.00)</td>
<td>3.74 (0.00)</td>
<td>1</td>
</tr>
<tr>
<td>R. IFG</td>
<td>1</td>
<td>0.19 (0.00)</td>
<td>3.74 (0.00)</td>
<td>4</td>
</tr>
<tr>
<td>B. OFC</td>
<td>4</td>
<td>0.14 (0.07)</td>
<td>3.36 (0.04)</td>
<td>5</td>
</tr>
</tbody>
</table>

Table 2: Differential effects of face stimuli as measured with gradient- and spin-echo EPI. N indicates the number of subjects who showed significant activation (p < 0.001, clusters of 4 or more contiguous voxels). Volumes and t-values were calculated before spatial normalization. Coordinates are in the normalized space of the Talairach and Tournoux brain atlas. For each region, mean volume, mean t-value and mean coordinates were averaged across all subjects. SEM are indicated in parentheses. L = left, R = right, M = medial.
DISCUSSION

In this study we directly compared fMRI activation measured with GE- and SE-EPI during visual perception of faces, which is mediated by neural activation within a distributed cortical network. With both fMRI techniques, bilateral activation was observed in multiple regions, including the IOG, FG, STS, amygdala, IFG, and the OFC. When compared with the GE sequence, activation measured with the SE sequence revealed smaller clusters with lower statistical significance, smaller amplitudes of the fMRI signal, and smaller number of subjects who showed the effect.

With both SE- and GE-based fMRI, we found that attentive viewing of faces evoked bilateral activation in the same regions of the distributed cortical network that mediates face recognition [14, 26]. These regions included the IOG and lateral FG, where face detection and identification is processed [27-29, 15, 30]; the STS, where information about social communication, such as the direction of the eye gaze, is processed [14, 31, 32]; and the amygdala and IFG, which mediate the perception of emotional facial expressions [33, 16, 34, 35]. We also observed activation in medial OFC, a region where face-selective neurons were found in the monkey brain [36]. It has been reported that patients with OFC lesions have difficulties identifying emotional facial expressions [37]. The OFC therefore seems to play an important role in the processing of facial cues that are important for social reinforcement [38]. Although all face stimuli evoked significant activation in all regions in both GE and SE.
experiments, we found that famous and emotional faces evoked stronger responses than unfamiliar faces, consistent with previous reports of valence enhancement [16, 39]. Finally, with both sequences we found hemispheric asymmetry with larger and more significant clusters of activation in the right than in the left hemisphere.

Previous studies have shown that inherent differences between GE and SE contrasts result in lower fMRI signal changes for the SE-based functional contrast [40, 13]. We compared activation evoked by faces in the visual cortex, limbic system, and prefrontal cortex, as measured with both GE and SE contrasts. While significant activation was found in the same regions in both experiments, the activation observed in the SE data was significantly reduced, as manifested by smaller and less significant clusters of activation and smaller amplitudes of the fMRI signal, consistent with previous studies [5, 41, 11].

SE sequences are less susceptible to magnetic field inhomogeneities. We therefore hypothesized that the MRI signal, and hence the functional signal, would be restored in regions that are prone to susceptibility-related signal dropout. We found that the MRI signal in the OFC and in the vicinity of the inner ear canals was substantially refocused in the SE-EPI images, while GE-EPI images exhibited severe signal void in these regions. Such signal dropout was caused by phase coherence loss in static field inhomogeneities, and was therefore considerably reduced by the 180° refocusing pulse of the SE sequence. Although signal in the OFC was restored, the inherently lower functional contrast of the SE sequence did not reveal stronger activation when compared with the GE sequence. Similar findings were reported by Gorno-Tempini et al. [42] in the anterior temporal lobes. A considerable amount of signal was restored by an echo-time reduction in a GE functional experiment, but the functional contrast at the lower echo-time was not sufficient to detect stronger activation. SE-based fMRI only performs better than GE-based fMRI in regions where its inherently lower functional contrast outweighs the signal (and hence functional signal) loss associated with the GE acquisition. This effect seems likely to occur in regions where signal is reduced by a factor of more than 2 in the GE as compared with the SE sequence. The reduction of functional contrast observed in our data is consistent with previous reports [5, 4, 13].

With both GE and SE sequences we detected activation in the same cortical regions. In contrast, a recent study comparing GE- and SE-EPI in a Stroop color-word matching task has reported activation in prefrontal cortex only with the SE sequence [4]. It is important to note, though, that with the lower spatial resolution and consequently larger voxel size used in that
study, signal in the OFC presumably suffered more from intravoxel dephasing when measured with GE-based fMRI. Additionally, the comparison between the two sequences was performed between two groups of subjects. We collected data at a higher spatial resolution and used parallel imaging that provided faster encoding and reduced echo-train length, and performed the comparison between the two sequences within the same subjects during the same experimental session. These factors can account for the different patterns of activation in OFC observed in the two studies.

In summary, significant activation within multiple, bilateral regions can be measured with both GE and SE contrasts. However, activation size and functional signal were significantly smaller in the SE data. In the OFC, a region prone to susceptibility-related signal dephasing, the SE acquisition has not revealed a significant advantage over the GE acquisition in terms of functional signal. Our data indicate that optimized GE sequences that reduce susceptibility artefacts are sufficient to detect activation in the OFC.

Acknowledgements

We thank Michaela Soellinger and Victor Candia for their helpful comments, Philips Medical systems for their continuous support, and the ETH Zurich program of Strategic Excellence Project (TH 7/02-02).

REFERENCES


COMPARISON OF "SILENT" CLUSTERED AND SPARSE TEMPORAL FMRI ACQUISITIONS IN TONAL AND SPEECH PERCEPTION TASKS

Conny F. Schmidt, Martin Meyer, Simon Baumann, Christof Baltes, Peter Boesiger, Lutz Jancke

under revision at NeuroImage

ABSTRACT

Sparse temporal acquisition (STA) schemes, which prevent interference between scanner noise and the auditory stimulus, have been developed and successfully applied in auditory fMRI. In order to combine this advantage with statistically more powerful data collection, we employed a clustered temporal acquisition (CTA). In the CTA three volume scans, as compared to only one volume scan in the STA, are acquired in rapid succession during each trial. The present study was carried out to directly compare the CTA with the STA scheme in the context of tone and auditory sentence perception. While previous studies employed the STA in short tonal stimuli and connected speech paradigms, the CTA approach has been used so far only in combination with short tonal stimuli. Irrespective of the acquisition technique, we localized activation in multiple, bilateral regions including Heschl’s gyrus, the planum temporale, the planum polare, and in the frontal opercular cortex even at the level of individuals. Consistently, average T-values were significantly higher in the CTA than in the STA data in all regions across the supratemporal plane. Our data demonstrate that the CTA scheme as compared to the STA approach provides enhanced statistical power per time unit and can be successfully applied to map auditory functions. Most importantly, we have shown
that CTA schemes are not restricted to tasks involving short tonal stimuli, but can reliably be applied to investigate the functional response to auditory sentences. It thereby opens new possibilities to investigate auditory sentence comprehension free of psycho-acoustic masking effects and attentional confounds.

INTRODUCTION

Functional magnetic resonance imaging (fMRI) has emerged as the most challenged technique for mapping human brain function and for the study of neuronal networks by probing spatial and temporal characteristics of the fMRI response. However, the acoustic noise generated by the fast switching gradients of echo-planar (EPI) sequences typically used for fMRI has particularly limited its application to audition and language comprehension research [1-4].

MR-related acoustic noise can directly confound mapping auditory functions at a perceptual level. Psychoacoustic masking occurs at a cochlear and cortical (foreground – background decomposition) level, when spectral components of the scanner noise overlap with the acoustic stimulus. At a cortical level, background scanner noise has shown to evoke an elevated baseline level of activation, which might considerably reduce the modulation range of the functional response to the target auditory stimulus (“clipping” [5]). Therefore, the auditory responses to the scanner noise and the target auditory stimulus might not add up linearly [6-8]. In addition to saturation effects on the neuronal firing rate or vascular dynamics, inhibition, habituation, attentional and distructional effects have been demonstrated to indirectly disrupt the auditory response (e.g. [4, 9, 10]).

The EPI acoustic noise is distributed over a complex frequency spectrum, and is manifested by noise bursts, which are generated when Lorentz forces act upon the gradient coils. Hence, physical features of the MR system have been engineered in order to reduce generation and transmission of acoustic noise. However, the majority of these hardware-related noise control solutions (for detailed review see [11, 3]) limit the benefit of high performance gradient capabilities. And admittedly, all these strategies are only capable of reducing but not annihilating ambient scanner noise.
Hence, fMRI designs have been adopted which present an auditory stimulus in silence intercept without interference with scanner noise (e.g. [1, 12, 6, 13, 7, 14, 8, 15-17]). These designs exploit the physiological delay between the onset of the stimulation and the succeeding hemodynamic response. Principally, the functional data are sampled within a time window close to the vertex of the hemodynamic response to the auditory stimulus at issue, and the acquisition window is chosen short enough to rule out that the response to the scanner noise interferes with the response to the stimulus. A subsequent, long silent period then allows both the response to the auditory stimulus and the response to the scanner noise to cease prior to the next trial. We would like to note that, in the context of our study, we relate the terms "sparse temporal acquisition" and "clustered temporal acquisition" to single and clustered VOLUMES of scans. Actually, our understanding of the notion "clustered acquisition" differs from the definition of Edmister and colleagues [7] who relate this expression to the clustered (as compared to the distributed) acquisition of SLICES WITHIN ONE VOLUME SCAN. The "clustered temporal acquisition" and the "sparse temporal acquisition" are from here on referred to as CTA and STA, respectively.

While sparse temporal sampling techniques have already been successfully employed in short tonal stimuli and connected speech paradigms (e.g. [18, 8, 19, 20]), the feasibility of the CTA - according to our definition - has so far only been demonstrated by two fMRI studies which presented short tonal stimuli [16, 17].

In the present study we conduct a systematic comparison between the CTA and STA scheme in two event-related auditory experiments. The CTA essentially combines a sparse temporal sampling scheme with a multiple volume acquisition. Specifically, we grouped all volumes of the time series into triplets with rapid intra-cluster succession, interspersed with long inter-cluster repetition times. We hypothesize that the data obtained with the CTA as compared to the STA would exhibit enhanced statistical power per time unit, according to the larger number of data acquired.

We applied the two acquisition schemes in two different auditory event-related experiments, during which the subjects were either presented with tonal stimuli or sentences. In accordance with previous studies we furthermore hypothesize that auditory tone and sentence comprehension recruits superior temporal areas in both the left and the right hemisphere and the bifrontal operculum [21-34].
More evidently, we predict more statistical power per time unit for the CTA as compared to the STA while subjects listen to both tones and sentences in a silent environment. We will demonstrate that, with the CTA, we employ an acquisition approach that avoids the detrimental influence of ambient scanner acoustic noise. Consequently it improves the subjects’ comfort, and concomitantly enhances the statistical power, allows the analysis of imaging data obtained from individuals and hence increases the validity of results.

MATERIALS AND METHODS

Subjects

Fifteen healthy volunteers (five males, ten females, age 27 ± SD 7 yr), all right-handed (14 ± 1 scores according to the Chapman & Chapman questionnaire [35]), participated in the study. All subjects were native speakers of German and have not been familiar with the stimuli prior to scanning. They had no neurological or aphasic history, nor did they have any hearing disorder. Written informed consent was obtained prior to the examination. The study was in accordance with Zurich Medical Faculty Ethical guidelines.

Stimuli and Design

The benefits and the reliability of the CTA as compared to the STA scheme were studied by performing two auditory experiments which comprised either tonal or auditory sentence stimuli. For each event class we collected two functional time series in randomized order with the CTA and STA (in total four functional time series) in the same 15 subjects.

Tonal Stimuli: In the tone experiment, stimuli consisted of sine wave, string, piano and trumpet tones, each 0.4 s long. Half the stimuli had a frequency of 349 Hz (F4, American notation) and half 466 Hz (B4, American notation). To maintain the subjects’ attention throughout the experiment, they were instructed to judge whether the tones were high or low pitch.

Speech Stimuli: In the speech experiment, stimulus material consisted of normal and pseudo sentences of 3.4 ± 0.4 s duration. In the normal speech condition subjects heard normal German sentences while the pseudo speech condition comprised grammatically correct
nonsense sentences with all content words replaced by phonotactically legal pseudo words (for details see [22]).

An exemplary sentence from the normal speech condition was:

\textit{Die besorgte Mutter sucht das weinende Kind.}

\textit{The anxious mother searches for the crying child.}

Likewise an example sentence from the pseudo speech condition was:

\textit{Das mumpfige Folofel hongert das apoldige Trekon.}

\textit{The mumpfy folofel hongers the apoldish trekon.}

In both normal and pseudo speech condition half the stimuli were active voice sentences and half were passive voice sentences. Throughout the second experiment the volunteers were to decide as to whether each sentence had an active or passive sentence structure. For technical reasons, however, recorded behavioral data got lost.

A total of 36 auditory events and 16 empty trials (silence) were presented in each run in randomized order to enable event-related analysis [36, 37]. In the STA, each trial started either at 3.5, 4.5, or 5.5 s prior to the data acquisition so that data from three distinct time points along the response were sampled [38]. In the CTA, however, no jittering was applied and each trial started at a fixed interval of 3.5 s before acquisition of first volume (see Figure 1).

All sound and speech files were digitized at a 16 bit/44.1 kHz sampling rate and were volume balanced using Volume Balancer software (Version 1.3, www.delback.co.uk/volbal/). Stimuli were controlled using Presentation® software (Version 0.70, www.neurobs.com). Stimulus presentation was synchronized by a 5 V TTL trigger pulse with the data acquisition. We used standard Philips headphones for binaural stimulus delivery.

\textit{Clustered sparse temporal acquisition}

We implemented a clustered sparse temporal acquisition (CTA) technique that combines the principle design of a sparse temporal acquisition (STA) with the clustered acquisition of three
consecutive volume scans per trial (Figure 1). Like the STA, the CTA scheme takes into account the slow, delayed onset of the hemodynamic response to any stimulus event. Previous studies showed that - similar to other auditory stimuli - acoustic scanner noise evokes a hemodynamic response within 3-5 s after the onset of the acoustic noise [2, 15]. Therefore, in the context of the present study the response to the stimulus does – if at all - only marginally interfere with the response to the scanner noise during an acquisition window of 3 s. Accordingly, our CTA scheme was designed with an acquisition time of 1 s for each volume and three consecutive volume scans. With the three volume scans the functional response to the stimulus was sampled at the plateau of the hemodynamic response with a delay of 3.5 s, 4.5 s and 5.5 s after stimulus onset [2, 12, 7, 8, 15].

![Figure 1: Sparse and clustered temporal acquisition (STA and CTA). With the STA the hemodynamic response (black line) is sampled only with one volume scan per trial, requiring three jittered events in order to obtain early and late responses, i.e. at 3.5 s, 4.5 s and 5.5 s after stimulus onset (dotted lines). In the CTA jittering of the events was not applied, as the response is obtained from three consecutive volume scans. Taken the hemodynamic response is slow and delayed (gray line), interference between the response to the stimulus and the response to the scanner acoustic noise is marginal.](image-url)
The rationale of a sparse (and in our study also of the clustered) temporal design is to present the auditory stimulus in silence, while the stimulus-evoked functional signal is acquired at the amplitude of the response within a short acquisition period. A long inter-scan interval (15 s in the STA and 12 s in the CTA, respectively in our study) then allows both the functional response to the auditory stimulus and the response evoked by the scanner noise to decay prior to the next trial. This approach is capable of separating the task-induced functional response from the scanner-noise induced functional response.

An exemplary signal time course obtained from the CTA functional time series is shown in Figure 2, and clearly revealed a significant response to speech stimuli as compared to silence. Signal intensity after stimulation was reliably higher than after empty trials and was consistently observed in all three timeframes in the cluster. The signal time course furthermore illustrates the decay of the MR signal due to $T_1$ saturation within each clustered acquisition ("decay-sampling").

![MR signal time course](image)

**Figure 2:** MR signal time course obtained from the CTA functional time series. Data were taken from a significant cluster of activation in left Heschl's gyrus in one subject. When compared to silence (empty trials), speech stimuli evoked a significant response. Signal intensity after stimulation (black) was reliably higher than after empty trials (grey) and was consistently observed in all three timeframes in the cluster. The time course furthermore illustrates the decay of the MR signal due to $T_1$ saturation within each clustered acquisition ("decay-sampling").
Data Acquisition

Measurements were performed on a Philips Intera 3 T whole body MR unit (Philips Medical Systems, Best, The Netherlands) equipped with an eight-channel Philips SENSE head coil. Functional time series were obtained from 14 transverse slices covering auditory cortex with a spatial resolution of 2.7 x 2.7 x 4 mm³ using a Sensitivity Encoded (SENSE [39]) single-shot gradient-echo planar sequence (acquisition matrix 80 x 80, SENSE acceleration factor R = 2.0, FOV = 220 mm, Tacq = 1000 ms, TE = 35 ms and θ = 90°). With the STA scheme, only one volume was acquired with a 15 s inter scan interval (ISI), while three volumes were acquired with the CTA scheme (Figure 1). Additionally we obtained one echo planar image that covered the whole brain with 38 transverse slices (Tacq = 4000 ms) but used otherwise the same scan parameters as for the functional time series. This whole-head EPI volume was used to assist the spatial normalization of the functional time series (see Data Analysis). Furthermore, we collected a standard 3D T₁ weighted scan for anatomical reference with 1 x 1 x 0.8 mm³ spatial resolution (acquisition matrix 224 x 224, TE = 2.30 ms, TR = 20 ms, θ = 20°).

Data Analysis

To account for different T₁ saturation effects in the CTA data, the three volume scans in each cluster were separated into three time series ("run1", "run2", "run3") during data analysis (Figure 3). Consequently, even if the longitudinal magnetization is partially saturated in the later volumes of the cluster, we contrasted the activation and baseline signals only from the "same-volume" time series. Furthermore, each of these three "runs" corresponds to the hemodynamic response sampled at a distinct temporal window, i.e. 3.5 s, 4.5 s and 5.5 s after stimulus onset. Post-processing and statistical analyses were carried out using SPM99 software (Wellcome Department of Cognitive Neurology, London, UK, www.fil.ion.ucl.ac.uk/spm/). All volumes were realigned to the first volume, corrected for motion artefacts, mean-adjusted by proportional scaling, normalized into standard stereotactic space (template provided by the Montreal Neurological Institute), and smoothed using an 8 mm full-width-at-half-maximum Gaussian kernel. Notably, in order to improve the normalization of the 14 slices obtained in each functional volume, the functional time series were first coregistered to the whole head EPI images. The whole head EPI images were then normalized into standard stereotactic space and the same transformation was applied to the
coregistered time series. The time series was furthermore high-pass filtered to eliminate low-frequency components (filter width 600 s). Statistical analysis was based on the General Linear Model [40, 41]. Single trials were treated as epochs and modelled by means of a box car function. We calculated contrast images by applying appropriate linear contrasts to the parameter estimates for the parametric regressor of each event vs. empty trials (Figure 3). Statistical parameteric (T) maps of these contrast images entered a one sample t-test in order to obtain the group main effects and were further subjected to a paired t-test in order to reveal significant differences in statistical sensitivity between the acquisition techniques (CTA vs. STA) at a voxel-by-voxel base. Clusters of activation were selected that showed a significant effect (p < 0.05, corrected for multiple comparisons). In the paired t-tests, the correction for multiple comparisons was restricted to the combined regions that showed significant activation in the group results in the CTA and STA acquisition.

Figure 3: Design matrices and first-level contrasts for a. the tonal and b. speech perception experiment. The functional time series obtained from the CTA was separated into three time series ("run1", "run2", "run3") according to the position of each volume scan within the acquisition cluster, in order to account for T<sub>1</sub> saturation effects. Similarly, different jitter positions in the STA, as indicated by 3.5 s, 4.5 s, and 5.5 s delay between stimulus presentation and acquisition, were modelled as separate eigenvectors. Significant clusters of activation were identified as the combined activation from all three "runs" in the CTA and all three jitter positions in the STA, respectively. Sin = sine wave tones, Tim = timbre tones (string, piano, trumpet). Norm = normal speech, Pseu = pseudo speech condition.
Region of Interest Analysis

In order to analyze differences in local brain activity and statistical detection power as a function of the acquisition technique (CTA, STA) and region (ROI), a set of six regions of interest was placed for each subject in the bihemispheric superior temporal plane (STP), namely in Heschl’s gyrus (HG, middle STP), the planum polare (PP, anterior STP) and planum temporale (PT, posterior STP). Regions of interest (Figure 4) were anatomically defined based on macroanatomical landmarks [42, 43]. The left and right HG ROI included 520 voxels centered at MNI coordinates \(x = +79, y = -15, z = 5\), the left and right PP ROI included 402 voxels with its center at \(x = -46, y = -4, z = -7\) and the PT ROI comprised 403 voxels at \(x = +54, y = -30, z = 12\). For each individual subject average T-values were obtained from each ROI and acquisition scheme, separately from the tonal and speech perception experiments. These mean values were subjected to a repeated-measure within-subjects effects (2 x 6) ANOVA with factors acquisition scheme (CTA, STA) and ROI (left and right PP, left and right HG, left and right PT). All main effects and interactions with two or more degrees of freedom in the numerator were adjusted with the procedure suggested by Huynh & Feldt [44]. The threshold for significance was set at \(p < 0.05\). Notably, ROI analysis based on T-values across anatomically predefined ROIs was mandatory, since a quantification on the basis of individual clusters is statistically difficult as not all subjects showed activation in all ROI and for both acquisition methods [45].

Figure 4: Location of ROIs placed in left and right Heschl’s gyrus (HG), in the planum polare (PP) and in the planum temporale (PT). L = left, R = right.
RESULTS

In order to localize clusters of activation that respond significantly to tonal and speech stimuli, we defined two contrasts, tonal stimuli versus silence (empty trials) and speech stimuli versus silence. Tonal stimuli were jointly comprised of sine wave, string, piano, and trumpet tones. Speech stimuli included both normal and pseudo sentences. We do not discuss differential brain responses evoked by sine wave and timbre tones or normal and pseudo speech, respectively.

Activation within a Network of Regions in Response to Tonal Stimuli

Perception of sine wave, string, piano, and trumpet tones as compared to silence, evoked a significant response in bilateral temporal and frontal regions (Figure 5), including Heschl's gyrus, the middle and superior temporal gyri (STG), insula and fronto-opercular region (FOR)

Figure 5: Activation evoked by tonal stimuli as measured with the CTA and STA. a. Shown from the left to the right are statistical (T) maps from four representative subjects, overlaid on transverse sections of each subject's normalized brain (Z = -2). Tonal stimuli as compared to empty trials evoked significant bilateral activation in Heschl's gyrus, in the planum polare and planum temporale (p < 0.05, corrected, min. 4 contiguous voxels). Consecutive rows represent data measured with the CTA and STA respectively. Subtraction of the T-value maps (only statistically significant clusters of activation) illustrates the differences in statistical sensitivity between the two acquisition techniques. b. Group results. Significant differences in the group T-maps were revealed in a paired t-test, shown in the bottom row (p < 0.05, corrected for multiple comparisons).
In the CTA data almost all subjects displayed significant activation in the STG and in the FOR. However, in the STA data only two third of the subjects displayed significant activation in the STG, and only one third of subjects additionally exposed significant activation in the FOR (Table 1). Consistently, clusters of activation were larger, and T-values were higher in the CTA as compared to the STA data. Multi-subject averaging shows the group effect. Paired t-tests performed between CTA and STA statistical parametric (T) maps revealed statistically more significant activation for the CTA bilaterally in superior temporal and in right deep fronto-opercular areas at the group level (Figure 5, Table 1). The STA did not reveal stronger statistical effect than the CTA in any region. Figure 7 further illustrates that in the CTA data a larger number of subjects showed significant activation in larger portions of the peri-sylvian cortices. Furthermore, the spatial extent of overlapping clusters of activation was considerably larger in the CTA as compared to the STA data.

Average T-values from left and right Heschl's gyrus (HG), planum polare (PP) and planum temporale (PT) entered an ROI analysis in order to elucidate potential local differences in the activation pattern. In almost all regions (except right PP) the mean T-values for the CTA were significantly higher as compared to the STA (p < 0.001). Specifically, T-values were about 50 - 90 % larger in HG and PT in the CTA than in the STA data (Figure 8), whereas the acquisition effect was less prominent in the PP (0 – 30 %). Furthermore, we found highest T-values in PT and lowest in PP, irrespective of the acquisition scheme.

A global (2 x 6) ANOVA with factors acquisition scheme (CTA, STA) x ROI (left and right PP, left and right HG, left and right PT) revealed a significant main effect of acquisition scheme (F_{1,14} = 7.87, p < 0.05), which reflects the higher statistical power (T-values) achieved by the CTA as compared to the STA data. Furthermore, a main effect of ROI (F_{5,10} = 5.37, p < 0.001) and an interaction acquisition scheme x ROI (F_{5,10} = 2.93, p < 0.05) demonstrated that the CTA differentially affected the T-values obtained from distinct ROIs. Thus, the finding that CTA least strongly affected the PP can be statistically underlined.
Perception and processing of auditory sentences as compared to silence, evoked a significant response bilaterally in temporal and frontal regions (Figure 6).

With both the CTA and STA we localized significant clusters of activation in all subjects bilaterally in the STG (Table 2). The CTA also uncovered fronto-opercular activation in 10 out of 14 subjects, while only 8 subjects displayed significant activation in the STA data. Notably, the size of clusters of activation and mean T-values were considerably larger for the CTA as compared to the STA. At the multi-subject average level paired t-test performed between STA and CTA statistical parametric (T) maps clearly revealed statistically more significant activation for the CTA in bilateral superior temporal areas. Furthermore, Figure 7 illustrates that a larger number of subjects showed significant activation in larger portions of the peri-sylvian cortices and that the spatial extent of overlapping clusters of activation was considerably larger in the CTA data. Intriguingly, the group effect plotted in Figure 6b does not unveil the enhanced statistical effect size in the FOR (higher T-values, larger clusters of activation) observed in the individual CTA data. Yet, Figure 7 depicts that the spatial extent
of overlapping clusters of activation in the IFG was considerably larger in the CTA as compared to the STA data.

![Figure 7](image_url)

**Figure 7:** Number of subjects that showed a significant response (p < 0.05, min 4 contiguous voxels) in a region, left: in the tonal perception task, right: in the speech perception task. Top row CTA, bottom row STA.

Analogous to the analysis of responses to tonal stimuli, average T-values taken from left and right HG, PP and PT entered an ROI analysis in order to elucidate potential local differences in the activation pattern. As apparent from Figure 8 in all regions the mean T-values for the CTA were significantly higher when compared to the STA (p < 0.001). Furthermore, we found highest T-values in HG and lowest in PP irrespective of the acquisition scheme.

A global (2 x 6) ANOVA with factors *acquisition scheme* (CTA, STA) x *ROI* (left and right PP, left and right HG, left and right PT) revealed a significant main effect of *acquisition scheme* ($F_{1,13} = 21.04$, $p < 0.001$). Thus, our finding that the CTA revealed higher statistical power (T-values) in the context of speech perception is supported by the ROI analysis. A significant main effect of *ROI* ($F_{3,9} = 19.22$, $p < 0.001$) and an interaction *acquisition scheme* x *ROI* ($F_{5,9} = 7.46$, $p < 0.001$) demonstrated that the CTA had a significantly different effect on the significance with which the functional response was detected in the distinct ROIs.
Figure 8: Comparison of CTA and STA performance in a tonal and speech perception experiment. Displayed are T-values, averaged across all subjects within anatomically defined ROIs in Heschl’s gyrus (HG), in the planum polare (PP) and planum temporale (PT). Error bars indicate the SEM. L = left, R = right.
### Main Effect of Tonal Stimuli

<table>
<thead>
<tr>
<th>Region</th>
<th>N</th>
<th>cluster size (k)</th>
<th>T-value</th>
<th>MNI coordinates</th>
<th>cluster size (k)</th>
<th>T-value</th>
<th>MNI coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Middle/ Superior temporal gyrus</strong></td>
<td>L</td>
<td>14</td>
<td>9.9 ± 0.8</td>
<td>-58 ± 2</td>
<td>7 ± 2</td>
<td>10</td>
<td>576 ± 166</td>
</tr>
<tr>
<td>R</td>
<td>15</td>
<td>1039 ± 199</td>
<td>9.9 ± 0.7</td>
<td>62 ± 2</td>
<td>-24 ± 3</td>
<td>-2 ± 2</td>
<td>11</td>
</tr>
<tr>
<td><strong>Insula, Fronto-opercular cortex</strong></td>
<td>L</td>
<td>12</td>
<td>7.6 ± 0.7</td>
<td>-39 ± 3</td>
<td>23 ± 3</td>
<td>2 ± 4</td>
<td>4</td>
</tr>
<tr>
<td>R</td>
<td>13</td>
<td>209 ± 81</td>
<td>6.8 ± 0.4</td>
<td>45 ± 2</td>
<td>24 ± 3</td>
<td>13 ± 5</td>
<td>5</td>
</tr>
</tbody>
</table>

Table 1: Local response maxima of clusters of activation for tonal stimuli as compared to silence (random-effects analysis, p < 0.05 corrected for multiple comparisons). Reported are T-values, cluster size k and location. Values were averaged over subjects (N) that showed a significant effect and listed as mean ± SEM. Regions that showed a significantly larger statistical effect size in the CTA as compared to the STA were identified in a paired t-test analysis. Where multiple foci existed in a cluster, the three most significant were listed. Coordinates are in the normalized space of the MNI brain. L = left. R = right. Labelling of anatomical regions was done by means of the Automated Anatomical Labelling toolbox implemented in SPM99 [46].
<table>
<thead>
<tr>
<th>Region</th>
<th>CTA</th>
<th>STA</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>N</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>cluster size</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>T-value</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>MNI coordinates</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>N</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>cluster size</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>T-value</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>MNI coordinates</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Region</th>
<th>CTA</th>
<th>STA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Middle/Superior temporal gyrus L 14</td>
<td>2930 ± 314</td>
<td>21.9 ± 2.4</td>
</tr>
<tr>
<td>Middle/Superior temporal gyrus R 14</td>
<td>3039 ± 362</td>
<td>21.3 ± 2.0</td>
</tr>
<tr>
<td>Insula, Fronto-opercular cortex L 10</td>
<td>421 ± 156</td>
<td>8.6 ± 1.0</td>
</tr>
<tr>
<td>Insula, Fronto-opercular cortex R 10</td>
<td>421 ± 130</td>
<td>9.6 ± 0.9</td>
</tr>
<tr>
<td><strong>CTA &gt; STA</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Middle/Superior temporal gyrus L</td>
<td>1343</td>
<td>7.72</td>
</tr>
<tr>
<td>Middle/Superior temporal gyrus R</td>
<td>1137</td>
<td>6.38</td>
</tr>
</tbody>
</table>

**Table 2:** Local response maxima of clusters of activation for speech stimuli as compared to silence (random-effects analysis, p < 0.05 corrected for multiple comparisons). Listing as in Table 1.

---

*The analysis of speech stimuli only includes data from 14 subjects, as one subject had to be excluded from the analysis due to motion artefacts in the part of the study during which speech stimuli were presented.*
**Temporal Dynamics of the Activation Pattern Evoked by Tones and Speech Stimuli**

In order to investigate the temporal dynamics of the activation pattern, we studied the T-values obtained from the statistical analysis of the three separate time series ("run1", "run2", "run3") of the CTA. These "runs" consist of the 1\textsuperscript{st}, 2\textsuperscript{nd} and 3\textsuperscript{rd} volume scans in the acquisition cluster, respectively, and thus represent three time series acquired 3.5 s, 4.5 s, and 5.5 s after stimulus onset. The functional response evoked by tonal stimuli revealed most significant activation in "run1", while T-values decreased significantly in "run2" and "run3" (p < 0.001, Figure 9). These data suggest either an underlying systematic effect (such as diminished signal-to-noise ratio in the late volume scans due to T\textsubscript{1} saturation or – unlikely – a habituational effect from acoustic scanner noise generated during the early volume scans), or reflect the underlying temporal dynamics of the functional signal (i.e., a short-lasting functional response). Since only 12 trials were acquired in each jitter position, the number of data was not sufficient to perform a similar analysis reliably with the STA data. We therefore only report the results from the CTA time series.

**Figure 9**: Mean T-values obtained from the statistical analysis of the three separate time series ("run1", "run2", "run3") from the CTA in the tonal and speech perception experiments. Each "run" represents a time series acquired 3.5 s, 4.5 s, and 5.5 s after stimulus onset (i.e., consists of the 1\textsuperscript{st}, 2\textsuperscript{nd} and 3\textsuperscript{rd} volume scans in the acquisition cluster). Mean T-values in the tonal perception experiment decrease over time after stimulus onset, while mean T-values increase and plateau in the speech perception experiment in all ROIs (HG, PP and PT).
The temporal dynamics of the functional response evoked by auditory speech as compared to silence, however, revealed that T-values were smallest in "run1", and plateaued during "run2" and "run3". This dynamic characteristic was consistently observed in all ROIs, and presumably reflects the processing complexity associated with sentences. It therefore represents an underlying physiological, rather than a systematic effect, which is in the context of a linear system that an extended stimulus duration evokes a prolonged hemodynamic response (Boynton et al., 1996).

Single Subject versus Group Analysis

Our data revealed that the effect of the acquisition scheme – CTA and STA – was particularly strong on the single subject level. The number of subjects that showed a significant response in a region was considerably higher in the CTA than in the STA data (Figure 7, Tables 1 & 2). Furthermore, consistently in all individual subjects, larger clusters of activation and higher T-values were observed with the CTA in all ROI (see single subject results Figures 5a, and 6a). While the multi-subject group analysis of the tone perception task clearly reflects this trend, the multi-subject group analysis of the speech perception task did not reproduce the individual differences (Figures 5b, 6b). The group statistical maps therefore imply that considerable inter-subject variance detrimentally limits the power of group statistics, although intra-subject statistics are improved.

DISCUSSION

In the present study we directly compared an STA with an CTA approach by means of two event-related auditory paradigms, that is perception of short tonal and auditory sentence stimuli. Consistently in each individual subject and within all ROIs activation measured with the CTA revealed larger clusters of activation and higher statistical significance as compared to the STA. Our data show that the CTA can be successfully applied to map auditory functions and provides enhanced statistical power per time unit. Most importantly, we have shown that the applicability of the CTA scheme is not restricted to tasks involving short tonal stimuli, but can successfully be applied to investigate auditory speech stimuli longer than one second or two.
We found that tonal stimuli evoked bilateral activation in superior temporal regions which have been formerly observed in the context of tone perception. As our tone stimuli comprised both sine wave and instrumental stimuli, it came at no surprise that we observed activation in cortical portions encompassing primary auditory cortex and which have formerly been attributed to the processing of complex auditory and musical timbre [47-49]. These regions sensitive to tonal stimuli are located on the temporal bank of the Sylvian fissure and cover Heschl’s gyrus which partly overlaps with the primary auditory cortex [50, 42], the anterior temporal plane (PP), and the posterior temporal plane (PT). The latter two regions have been associated with complex auditory functions and can be considered part of the auditory cortex [24, 51, 27, 30, 52-54]. Notably, tone perception also evoked significant activation in frontal portions of the peri-sylvian region in almost all subjects, namely in the deep frontal operculum lining the anterior insula. Even though these regions do not exclusively subserve auditory perception, this deeply buried part of the frontal operculum has been observed in neuroimaging studies on auditory perception when subjects had to perform demanding tasks, i.e. during imaging timbre [47], categorical perception [55], sound categorization [56], pitch memory [57], and pitch discrimination tasks.

Auditory sentences evoked significant activation bilaterally along the peri-sylvian cortex, consistent with present knowledge on the cerebral organization of speech perception [58, 59, 22, 60-63]. Our speech stimuli also evoked a significant response in the frontal part of the peri-sylvian region. However, the center of gravity was localized more laterally (Tables 1 & 2, Figure 6), close to the convexity of the inferior frontal gyrus (IFG). Presumably, it appears that interindividual variance of frontal activation may account for this finding [64, 65]. Alternatively, speech perception may involve both the deep frontal operculum and the lateral convexity of the IFG [66].

In the present study we conducted a systematic comparison between the CTA and STA scheme in two event-related auditory experiments, and aimed to elucidate the statistical benefits of the CTA approach and its practical implications. Essentially, in event-related sparse temporal designs the auditory stimulus is presented during silent inter-scan-intervals which are not contaminated by scanner noise. Consequently, the number of images per time unit is considerably smaller in sparse temporal schemes as compared to continuous scanning, and inter-scan-intervals of relatively long duration are required to uncouple the hemodynamic responses from succeeding transient single events, extending the total scan time remarkably.
Longer imaging time, in turn, might be paralleled by attention loss and subsequently lower functional response [9]. However, several studies have shown that the functional response to an auditory stimulus measured with a silent sequence as compared to continuous scanning was enhanced in terms of the amplitude of the response and spatial extent of the activation [16]. Most importantly, it has been demonstrated that sparse temporal acquisition schemes ensure that only the acoustic stimulus contributes to the functional response, resulting in enhanced spatial and functional specificity of the response [8, 16]. In order to combine the advantage of silent event-related sparse temporal designs with efficient data collection, we used a clustered temporal acquisition procedure. CTA schemes with long acquisition duration have already been successfully employed to study the functional response to short sine wave tones [6], to EPI noise itself [2], and in motor tasks [2, 13]. In these latter studies, late volume scans within the cluster were progressively contaminated by the functional response evoked by the scanner noise generated during the "early" volume scans. Even very recently, clustered acquisition schemes with only short acquisition duration have been applied in order to avoid this intra-acquisition interference. However these CTA approaches have so far been used only in the context of short tonal stimuli [16, 17]. In the present study we therefore aimed to compare the CTA and STA in the context of tone perception and auditory sentence processing, which involved stimuli of extended duration.

During the clustered acquisition the longitudinal magnetization of consecutive volumes within the acquisition cluster was partially saturated due to $T_1$ relaxation. This particular kind of design is also called a "decay-sampling scheme", since the functional signal is acquired during non-steady-state conditions of the longitudinal magnetization. Therefore, we contrasted stimulation and empty trials only from "same volumes" in the time series ("run1", "run2" and "run3"). We assume that the magnetization saturation is equal during activation and baseline, and attribute any signal intensity difference to a functionally driven signal change. In order to separate saturation effects from functional signal changes, the EPI acquisition can be driven into equilibrium with preceding silent saturation pulses [67]. However, most decay-sampling studies simply contrasted scans with and without the target stimulus under the assumption that $T_1$ decay is the same during activation and baseline trials.

Under the simplifying assumption of independent "runs" we analyzed the three consecutive volumes of each cluster as a separate time series. Since significant clusters of activation were identified from the combined activation from all three "runs" in the CTA, we hypothesized a
\sqrt{3} enhancement of the contrast to noise ratio and therefore a concomitant increase in statistical power per time unit. Notably, the \sqrt{3} relationship holds only true in case of a real underlying effect. Indeed, we found statistical power increased by a factor close to the theoretical value in HG and in the PT, whereas the PP received only a smaller benefit from the CTA. The findings are in agreement with our predictions since the PP showed only a small functional effect (low average T-values) in response to both tonal and speech stimuli. That is, our data approached the theoretical, simplified estimation of \sqrt{3} the higher the underlying average T-values within the region were. We conducted a repeated-measure (2 x 6) ANOVA in order to test for regional differences in activation pattern and statistical detection power as a function of acquisition scheme and region. Consistent with the findings described above, we observed a significant main effect of acquisition and a significant acquisition x ROI interaction which clearly demonstrates the statistical effect size (mean T-values) was significantly larger in the CTA data, and that not all regions received the benefit to the same extent from the CTA as compared to the STA.

In order to study the temporal dynamics of the activation pattern and possible interactions between the three separate CTA time series ("run1", "run2" and "run3"), we compared the T-values obtained separately from these three runs. While the T-values revealed from the tonal perception task decrease from the early to late volume scans, they plateau in the speech perception experiment. These results suggest that neither decay sampling nor a potential confounding habituation effect in the "late" volume scans were the major determinants for the statistical effect size obtained from the three separate CTA time series. Consistently decreasing - rather than increasing - T-values should then have been obtained between "run1", "run2" and "run3" in both the tonal and speech perception task. Instead, the temporal progression of statistical effect size (T-values) in our data was consistent with the temporal dynamics of the functional response to short and prolonged stimulus duration. Consistent with the longer stimulus duration, the functional response to auditory sentences was more prominent than the response to short tonal stimuli, both in terms of significance and spatial extent of the activation. Although we can not completely rule out a possible interaction between scanner noise and functional response to the auditory stimulus in "late" volume scans of the cluster (i.e. in "run2", "run3"), these data showed that even if such an interaction would exist, it does not significantly affect our results. Moreover strong evidence stemming from previous literature indicates that the onset of the functional response to an auditory event only starts at about 3-5 s after stimulus onset [2, 68, 15]. This suggests that in our study late
volume scans were only marginally – if at all - influenced by the functional response to the scanner noise of earlier volumes within the same cluster.

The present study primarily focused on the enhanced statistical effect size associated with the CTA as compared to the STA. We observed larger clusters of activation, and average T-values were significantly higher in the CTA than in the STA data. However, this effect was most prominent on a single subject level as well as in a multi-subject ROI analysis (where local differences were averaged across a whole region), while the multi-subject group analyses only partially mirrored this result: The multi-subject group analysis of tonal perception clearly reflects this trend, however, the multi-subject group analysis of the speech perception task did not reproduce the individual differences. Since tonal stimuli require only a relatively simple level of auditory perception, the evoked functional response pattern is comparatively consistent between individuals. However, during speech comprehension, multiple phonological, semantic, syntactic, and prosodic processing levels are involved and therefore the functional responses might per se be more variable in strength and spatial distribution between subjects. This may account for the distinct reflection of the single-subject results in the voxel-by-voxel group maps. Our data demonstrate that activation foci in the speech condition appeared highly variable between subjects resulting from inter-individual functional and anatomical variations although standard normalization to MNI coordinate system had been applied [65, 69, 50, 70]. This finding further contributes to the often-raised question on how anatomical localization and assignment can be performed in the presence of large individual anatomical and functional variability [71] and to what extent a voxel-by-voxel multi-subject group analysis is capable of portraying the whole variety of individual response patterns. At least we think it is recommendable to carefully study individual contrast images prior to multi-subject averaging.

In summary, we have demonstrated that the CTA technique provides a valuable and effective tool, which is particularly tailored to the specific demands of auditory fMRI and at the same time provides enhanced sampling efficiency. Hence it might help increase our knowledge on the functional architecture of auditory perception and to identify neural pathways subserving specific auditory and language functions in a silent scanning environment. We believe that these amendments help improve the applicability of event-related functional MRI in clinical settings, as in particular patients may benefit from this comfortable scanning procedure which nonetheless provides sufficient statistical power to obtain powerful statistical maps from
individuals. Finally, this novel approach is evidently applicable in studies examining the level of spoken stimuli longer than one word or two, which opens a new horizon in the realm of neurocognitive research of speech and language.

CONCLUSION

We employed a clustered and a sparse temporal acquisition in two event-related auditory tasks, which used either tonal or auditory sentence stimuli. The CTA and STA scheme allowed for silent fMRI scanning without interference of scanner noise, and thus the two approaches provide important tools for auditory fMRI devoid of psychoacoustic masking effects. We localized activation in regions known to mediate auditory perception and language comprehension and demonstrated the clear advantages of the CTA as compared to the STA. With the CTA we found larger clusters of activation consistently in all subjects and average T-values were significantly higher within all regions of interest. Therefore our data show that the CTA allows for time-efficient data collection which is paralleled by enhanced statistical power per time unit. This result implicates, that the gain in statistical power provided by the clustered acquisition might conversely be sufficient to reduce the number of trials required to perform a reliable statistical analysis, and therefore a reduction of total scan time.

Acknowledgments

This work was supported by a grant from the Swiss National Science Foundation 'Short-term and long-term plasticity in the auditory system' (46234101). The authors also wish to thank for the continuing support of Philips Medical systems and the financial support by ETH Zurich (Strategic Excellence Project TH 7/02-02).
REFERENCES


ABSTRACT

Sparse temporal acquisition schemes have been adopted to investigate the neural correlates of human audition using blood-oxygen-level dependent (BOLD) based functional magnetic resonance imaging (fMRI) devoid of ambient confounding acoustic scanner noise. These schemes have previously been extended to clustered-sparse temporal acquisition designs which record several subsequent BOLD contrast images in rapid succession, in order to enhance temporal sampling efficiency. In the present study we demonstrate that an event-related task design can effectively be combined with a clustered temporal acquisition technique in an auditory language comprehension task. The same fifteen volunteers performed two separate auditory runs which either applied customary fMRI acquisition (CA) composed of continuous scanner noise or “silent” fMRI built on a clustered temporal acquisition (CTA) protocol. In accord with our hypothesis, the CTA scheme relative to the CA protocol was accompanied by significantly stronger functional responses along the entire superior temporal plane. By contrast, the bilateral insulae and anterior cingulum engaged more strongly during continuous scanning which may be associated with unspecific nonauditory processes provoked by the noisy environment. A post-hoc region-of-interest
analysis revealed functional activation in subportions of the supratemporal plane which varies as a function of acquisition protocol. Evidently, the ambient scanner noise compromises auditory processing in the right planum polare and Heschl’s gyrus while the planum temporale exposes a leftward asymmetry irrespective of acquisition technique. Our findings therefore implicate that silent fMRI is advantageous should fMRI studies unambiguously explore auditory and speech functions residing in the supratemporal plane.

INTRODUCTION

Since the advent of neuroimaging techniques approximately twenty years ago our understanding of the cerebral organization of speech functions has substantially changed and magnified. However, customary functional magnetic resonance imaging (fMRI) is hampered by acoustic noise produced during operation of the MR system, which might affect the experimental integrity of auditory fMRI studies [1-3]. In particular, the MR acquisition inevitably produces bursts of acoustic noise up to 120 dB SPL which obscures the quality of auditory stimulus presentation. At the perceptional level recognition of spoken utterances is distorted by acoustic screening of the stimulus when its spectral components overlap with the scanner acoustic noise as well as by psycho-acoustic effects, such as a stapedal muscle reflex that alters the perceived sound level and spectral characteristics of the presented stimulus. At the physiological level the ambient scanner noise is tied to a saturation of the neuronal population residing in auditory cortices in the supratemporal plane. At the cognitive level the extraction of an auditory stimulus in an obscuring acoustic background is affected by foreground-background decomposition processes and attentional effort which may provoke additional (unwanted) functional responses in extra-auditory frontal areas. Hard-ware based attempts to alleviate scanner noise at its very source, e.g. engineering of Lorentz force-balanced gradient coils, election of gradient coil material with high stiffness in order to minimize mechanical vibration, implementation of acoustically damped mounting systems as well as special sound insulation have been successfully applied [4-6, 3]. In addition, recent evaluation of one experimental design to reduce scanner noise, the sparse temporal acquisition technique (STA) demonstrates the advantages of this approach in auditory experiments [1, 7-17]. Using this technique, the interval between single volume scans is enhanced and the physiological delay between the onset of the stimulation and the succeeding hemodynamic response is used to separate the functional responses evoked by the scanner.
acoustic noise and the auditory stimulus at issue. Thus, STA schemes allow insertion of silent intervals during which auditory stimuli are presented. As functional images are collected only instantly after stimulus presentation, the effective presentation of auditory events is not masked by the scanner noise and hence the aforementioned detriments pertaining to the perceptual, physiological, and cognitive level could be annihilated. To investigate the cerebral substrates of auditory speech perception it has become convenient to use scanning protocols which are reliant on the sparse sampling technique to avoid any interference with ambient scanner noise. Evidently, the functional response to an auditory stimulus as measured with a STA was magnified in terms of amplitude of the response and spatial extent of significant clusters of activation. Moreover, it has been shown that functional activation was localized with enhanced spatial and functional specificity [17]. However, compared to continuous scanning these acquisition schemes require longer total imaging time in order to collect sufficient data for powerful statistical analyses. Hence, “silent” event-related sparse temporal acquisition designs have been further elaborated upon to produce clustered-sparse temporal acquisitions (CTA)\(^5\). In this acquisition design multiple volume scans are acquired in rapid succession after each trial (therefore "clustered"), in order to combine advantages of "silent" scanning with time-efficient data collection. Incorporating long acquisition intervals, such schemes have been employed in a variety of studies, such as to investigate functional activation evoked by scanner noise itself [18], motor activation [18, 8], and in the context of short sine wave tone perception [19] as well as in fMRI-studies on syllable processing [20], auditory imagery [21], and short sentence stimuli [22]. Like the present study the paper by Rodd and colleagues also researched sentence comprehension but did not explicitly compare CTA vs. CA to reveal potential benefits and detriments tied to either one or the other approach.

The present study

The current study systematically investigates to what extent ambient scanner noise bound to continuous fMRI may affect hemodynamic responses to spoken sentences. As it has formerly been shown, scanner noise not only produces activation in brain regions recruited in audition

\(^5\) Importantly, we would like to emphasize that in the context of our study we relate the term "clustered temporal acquisition" to clustered volumes of scans, as opposed to the clustered acquisition of slices within one volume scan [9]. The clustered sparse temporal acquisition and the continuous acquisition techniques are from here on referred to as CTA and CA techniques, respectively.
but also interferes with stimulus presentation [1]. We therefore hypothesize that in the presence of continuous scanner noise the functional response to an auditory stimulus might be reduced in primary and secondary auditory areas. This prediction is predicated on the observations that persistent scanner noise leads to a steady saturation of neural ensembles in the auditory cortex and hence yields a reduced functional response to an additional auditory stimulus. Notably, a noisy environment raises the BOLD baseline in brain regions tied to auditory stimulation, and hence reduces the functional modulation range within the cortex can respond to an auditory stimulus. Conversely, we expect an enhanced BOLD response for the silent CTA, since this silent acquisition design produces less BOLD baseline activation and hence stronger functional responses to sound stimuli. It is of interest that BOLD baseline activation is enhanced in an acquisition pertaining pulsed gradient sounds, but can be reduced in a neurophysiologically tuned continuous-sound fMRI sequence [23]. Interestingly, recent MEG research has demonstrated a “noise suppression effect” during audition. This study provided evidence that simultaneous presentation of spoken sentences and white noise particularly compromises auditory processing in the right hemisphere [24]. Hence, the present study should uncover to what extent scanner noise may persistently affect regional functional activation in auditory fields of the left and right hemisphere. Furthermore we assume that behavioural performance should be slower and possibly less accurate during “noisy” CA relative to “silent” CTA as the latter is presumed to provide facilitated auditory perception.

MATERIALS AND METHODS

Subjects

Fifteen healthy volunteers (8 males, 7 females, age 26 ± SD 5 years), all consistently right-handed as determined by the Annett-Handedness-Questionnaire [25, 26], participated in the study. All participants were native speakers of German and have not been familiar with the stimuli prior to scanning. They had no neurological or psychiatric history, nor did they have any hearing disorder. Written informed consent was obtained prior to the examination. The study was in accordance with Zurich Medical Faculty Ethical guidelines.
Stimuli and Design

Stimulus material consisted of normal and pseudo sentences of 3.4 ± 0.4 s duration. In the normal speech condition participants heard normal German sentences while the pseudo speech condition was composed of grammatically correct nonsense sentences with all content words having been replaced by pseudowords adhering to the phonotactical rules of the German language (for details see [27]).

An exemplary sentence from the normal speech condition was:

*Die besorgte Mutter sucht das weinende Kind.*

*The anxious mother searches for the crying child.*

Likewise an example sentence from the pseudo speech condition was:

*Das mumpfige Folofel hongert das apoldige Trekon.*

*The mumpfy folofel hongers the apoldish trekon.*

In both normal and pseudo speech condition half the stimuli were active voice sentences and half were passive voice sentences. We controlled all normal and pseudo sentences for syntax, mean duration, and mean amplitude.

All sound files were digitized at a 16 bit/44.1 kHz sampling rate and were volume balanced using Volume Balancer software (Version 1.3, www.delback.co.uk/volbal/).

Task

Throughout the experiment participants were asked to indicate as to whether each sentence had an active or passive syntactic structure and gave their response by button press (only using the right hand) as soon as they identified the sentence structure. During empty trials participants were also requested to give a button response. Prior to scanning participants underwent a brief training during which they were presented with spoken normal and pseudo sentences specifically recorded for practice purposes.
Procedure

During scanning the room lights were dimmed and a fixation cross was projected via a forward projection system onto a translucent screen placed at the end of the magnet’s gurney. Subjects viewed the screen through a double mirror attached to the head coil. For subject comfort the inner bore of the magnet was isolated with acoustic damping material (TRANSONIC, Spectra, Richwiller, France) in order to reduce acoustic noise transmission. Additionally, subjects were placed on an acoustic damping mattress in order to further reduce transmission of scanner acoustic noise via bone conduction. Stimuli were controlled using Presentation® software (Version 0.70, www.neurobs.com). Stimulus presentation was synchronized with the data acquisition by a 5 V TTL trigger pulse. We used an MR-compatible piezoelectric auditory stimulation system incorporated into standard Philips headphones for binaural stimulus delivery, enabling the transmission of strong sound pressure levels with excellent attenuation characteristics. Each volunteer performed two experimental runs (CA and CTA). We balanced the order of acquisition scheme across the subjects. A total of 60 auditory events and 20 empty trials (silence) was presented in each experimental run in an event-related pseudo-randomized order. Each sentence was only presented once either during CA or CTA. Trial duration was 15 s and the stimulus was presented 4.0 or 4.5 s prior to the data acquisition in the CTA. In the CA we varied the inter-stimulus-interval systematically by moving the onset of the stimulation in three steps of 0.5 s (cf. Figure 1) [7, 28]. Each trial was initialized visually by a 2 s fixation cross, which directly preceded the auditory stimulus or empty trial.

Data Acquisition

Measurements were performed on a Philips Achieva 3 T whole body MR unit (Philips Medical Systems, Best, The Netherlands) equipped with an eight-channel Philips SENSE head coil. Functional data were obtained from 14 transverse slices covering the entire peri-sylvian cortex parallel to the AC-PC plane with high spatial resolution of 1.7 x 1.7 x 3 mm³ using a Sensitivity Encoded (SENSE, [29]) single-shot gradient-echo planar sequence (acquisition matrix 128 x 128, SENSE acceleration factor R = 2.7, FOV = 220 mm, inter-slice gap 0.3 mm, TE = 35 ms).

With the CTA scheme, three subsequent volumes were acquired per trial with a TR 1 s, flip angle 90° (decay sampling), and an 12 s inter scan interval (ISI) resulting in 240 dynamic
scan volumes. With the CA scheme, ten volumes were acquired per trail continuously with a TR 1.5 s, flip angle 72° (total of 800 dynamic scan volumes). Three dummy scans preceded the experiment to allow the magnetization to reach a steady state. For both acquisition schemes—CA and CTA—the total scan time was kept constant.

![Clustered sparse acquisition diagram](image)

**Figure 1**: Clustered temporal and continuous acquisition (CTA and CA). With the CTA the auditory stimulus was presented during silence, followed by three volume scans (each with a TR = 1 s). Trial duration was 15 s with an 12 inter-scan interval (ISI). Stimulus presentation started 4.0 s or 4.5 s (jitter, dotted lines) prior to the data acquisition, forming a 0.5 s jitter interval. The stimulus induced functional signal (thin black line) was sampled at the presumed amplitude of the response and was separated from the scanner-noise induced functional response (thick grey line). During continuous data acquisition volume scans were acquired with a TR = 1.5 s. The stimulus presentation was jittered by 0 s, 0.5 s and 1 s respectively, forming a 0.5 jitter interval.

We additionally obtained one single volume scan that covered the whole brain with 38 transverse slices (TR 3000 ms, flip angle 82°) but otherwise the same scan parameters as used for the functional time series. This whole-head EPI volume was used to assist the spatial normalization of the functional data (see Data Analysis). Furthermore, we collected a standard
3D T₁-weighted scan for anatomical reference with 1 x 1 x 0.8 mm³ spatial resolution (acquisition matrix 224 x 224, TE 2.30 ms, TR 20 ms, flip angle 20°).

Data Analysis

Data post-processing and statistical analyses were carried out using SPM99 software (Wellcome Department of Cognitive Neurology, London, UK, www.fil.ion.ucl.ac.uk/spm/). All volumes were realigned to the first volume, corrected for motion artifacts, normalized into standard stereotactic space (voxel size 2 x 2 x 2 mm³, template provided by the Montreal Neurological Institute), and smoothed using a Gaussian kernel size 8 mm full-width-at-half-maximum. Statistical analysis was based on the General Linear Model [30]. The time series was high-pass filtered to eliminate low-frequency components (128 s in the CA) and mean-adjusted by proportional scaling. Single trials of the CTA time series were treated as epochs and modeled by means of a boxcar function. In order to account for T₁-decay effects in the CTA time series, we included three regressors into the GLM that take into account different MR signal intensities along the T₁-decay curve. The event-related functional response as measured by the CA was analyzed using a linear convolution model with an assumed hemodynamic response function [30, 31]. We calculated contrast images by applying appropriate linear contrasts to the parameter estimates of each event versus empty trials. These contrast images entered a one sample T-test in order to obtain the group main effects and were further subjected to a paired T-test in order to reveal significant differences between the acquisition techniques (CTA and CA) on a voxel-by-voxel basis. Clusters of activation were selected that showed a significant effect (p < 0.05 at cluster level, corrected for multiple comparisons).

Region of Interest Analysis

In order to analyze differences of local brain activity in auditory fields stretching along the entire supratemporal plane (STP) as a function of the acquisition technique, region and hemisphere, a set of six distinct regions of interest (ROI) was placed for each subject in Heschl’s gyrus (HG, middle STP), the planum polare (PP, anterior STP) and planum temporale (PT, posterior STP).

Regions of interest were anatomically defined based on macroanatomical landmarks [32, 33]. The left and right HG ROI included 520 voxels centered at MNI coordinates x = +/-49, y = -
15, z = 5, the left and right PP ROI included 402 voxels with its center at x = +/-46, y = -4, z = -7, and the PT ROI comprised 403 voxels at x = +/-54, y = -30, z = 12 (Figure 2). Mean t-values were collected for each subject from each ROI (HG, PP, PT), acquisition technique (CTA and CA), and sentence type (normal and pseudo speech) and were subjected to a repeated-measure within-subjects (2 x 3 x 2) ANOVA with factors acquisition scheme (CTA, CA), ROI (HG, PP, PT), and hemisphere (left, right). Due to different scaling of β-values obtained from different models (event-related analysis for CA and epoch-related analysis for CTA) we took t-values as proper effect sizes to compare the CA and CTA techniques. Calculating effect sizes based on t-values (or Z-values respectively) is an established approach in statistical analyses of fMRI data [34, 35]. All main effects or interactions with two or more degrees of freedom in the numerator were adjusted with the procedure suggested by [36]. The threshold for significance was set at p < 0.05.

![Figure 2: Size and position of three anatomically defined regions of interest in the bilateral superior temporal plane. HG = Heschl’s gyrus; PP = planum polare; PT = planum temporale. For more details see Methods section.](image)

RESULTS

Behavioural Data

The mean accuracies with which the subjects identified active and passive sentence structures in the presented auditory sentence material were 98% and did not differ between normal and pseudo sentences or CTA and CA, respectively. We did not find any significant differences in accuracy. A 2 x 2 ANOVA with factors acquisition (CTA, CA) x sentence type (normal, pseudo) uncovered that mean reaction times were significantly faster during "silent"
clustered (RT = 2.98 ± SD 0.73 s) than during continuous scanning (RT = 3.22 ± SD 0.73 s, F = 14.01, p > 0.005). We furthermore observed that volunteers responded significantly faster to normal (RT = 3.02 ± SD 0.18 s CA, RT = 2.79 ± SD 0.18 s CTA) relative to pseudo sentences (RT = 3.42 ± SD 0.19 s CA, RT = 3.17 ± SD 0.19 s CTA) irrespective of acquisition scheme (F = 106.05, p < 0.001).

*Imaging data*

Akin to former studies using the same stimulus corpus we [27, 37, 38] collected significant functional activation in response to auditory speech stimuli as compared to silence from bilateral peri-sylvian regions, in particular along the supratemporal plane, including Heschl's gyrus (HG), the plana temporale (PT) and polare (PP), and from the deep fronto-opercular region (FOR) encroaching onto the insula. Multi-subject data revealed a dissociation between frontal and temporal sites as a function of the acquisition scheme. Whilst we observed more significant activation bilaterally in superior temporal cortex (and the left lateral convexity of the pars triangularis) with the CTA, we noted stronger involvement of deeply buried insulae and anterior cingulum with the CA (Figure 3, Table 1).

![Figure 3](image.jpg)

*Figure 3:* Activation evoked by auditory sentence stimuli as measured with the CTA compared to CA overlaid on five transverse sections of the group average normalized brain. A paired T-test was performed on the CTA and CA single subject's contrasts, elucidating regions that were found significantly more (yellowish) activated for CTA or more strongly recruited during CA (bluish). The level of significance was set at p < 0.05 at cluster level, corrected for multiple comparisons (L = left, R = right).
Table 1: Local response maxima of clusters of activation that either showed a significantly larger statistical effect size with CTA (left column) or CA (right column) (random-effects analysis, p < 0.05 corrected for multiple comparisons) as revealed by a paired T-test analysis over the whole subject group (random-effects analysis, p < 0.05 at cluster level, corrected for multiple comparisons). The table reports t-values, cluster size k and location. Where multiple foci existed in a cluster, the three most significant were listed. Coordinates are in the normalized space of the MNI brain. L = left. R = right. Labelling of anatomical regions was done by means of the Automated Anatomical Labelling toolbox implemented in SPM99 [39].

Region of Interest Analysis

We computed a ROI analysis to reveal local differences and interaction effects between acquisition scheme, ROI, and hemisphere in the supratemporal response pattern. As apparent from Figure 4 we found consistently larger functional effect size with the CTA relative to the CA in all three supratemporal ROIs irrespective of hemisphere.

A global 2 x 3 x 2 ANOVA with factors acquisition scheme (CTA, CA), ROI (HG, PP, PT), and hemisphere (left, right) revealed a significant main effect of acquisition scheme (F = 61.55, p < 0.001), indicating a general stronger response for CTA, ROI (F = 13.11, p < 0.001), and hemisphere (F = 8.76, p < 0.01). Furthermore we identified a significant interaction between acquisition scheme and ROI (F = 21.60, p < 0.001), an interaction between ROI and hemisphere (F = 12.75, p < 0.001), and of acquisition scheme by ROI by
hemisphere (F = 8.25, p < 0.005), indicating that the functional response is modulated differently in distinct regions.

Figure 4: ROI analysis in the superior temporal plane elucidates local differences in the activation pattern obtained in the CTA and CA data for the main effect of auditory sentence stimuli as compared to silence. Effect size averaged across all subjects within anatomically defined regions of interest in Heschl’s gyrus (HG), in the planum polare (PP) and planum temporale (PT). Error bars indicate the SEM. L = left, R = right. * denotes significant hemispheric differences (p < 0.05).

Differences in mean effect size indicate that CTA corresponds to a relatively stronger increase of functional effect size in the right hemisphere in two out of three ROIs (PP, HG). To test the statistical significance of this regional difference in functional lateralization we subjected the data to a 2 x 2 ANOVAs with factors acquisition technique (CA, CTA) and hemisphere (left, right) for each ROI separately. The pattern of results for PP and HG is highly similar while it turns out differently for the PT. The 2 x 2 ANOVA for the PP evinced a main effect of acquisition technique (F = 5.18, p < 0.05), no main effect of hemisphere, but an interaction between acquisition type and hemisphere (F = 6.55, p < 0.05). The 2 x 2 ANOVA for the HG uncovered a main effect of acquisition technique (F = 69.67, p < 0.001), no main effect of hemisphere, but an interaction of acquisition type with hemisphere (F = 8.28, p < 0.01). In particular, this interaction of acquisition type and hemisphere we found for the PP and the HG implies that the effect size was stronger in the right relative to the left hemisphere with the “silent” CTA approach. For the PT the analysis yields a different pattern. A 2 x 2 ANOVA for the PT unveiled a main effect acquisition technique (F = 28.49, p < 0.001) and a main effect of hemisphere (F = 20.23, p < 0.001), but no interactions. As apparent from Figure 4 the PT exhibited consistently larger effect size for the CTA as for the CA schema and...
explored a generally stronger involvement of the left compared with the right PT. Since separate ANOVAs for the PP and the HG did not reveal main effects of hemisphere we conclude that the ambient scanner noise differentially affects the posterior STP (PT) and the middle and anterior portions of the STP (PP, HG).

Taken together, the fMRI results indicate that “silent” CTA generally corresponds to a more salient functional responses in auditory fields and that particularly the right auditory cortex (PP, HG) appears to be detrimentally affected by a noisy scanning environment. The CA, however, appears to lead to stronger brain responses in bilateral non-auditory frontal regions, namely the insula and the anterior cingulum.

DISCUSSION

In this study we compared a clustered “silent” acquisition with continuous an fMRI acquisition in order to test the potential advantage of the fMRI in “silent” environment in the context of an auditory speech experiment. Ever since functional MRI has been established as one of the major tools of Cognitive Neuroscience the deleterious effects of ambient scanner noise on auditory cortex functions have been debated [1, 18, 40, 2, 3, 41-43]. To overcome this drawback, an experimental design to reduce scanner noise, the sparse temporal sampling technique has been developed and successfully applied in auditory experiments [44, 8-10, 45, 11, 12, 46, 13, 15, 17]. As a recent advancement of the sparse technique clustered imaging comprising multiple rapidly succeeding volumes has been applied in the context of fMRI studies on sound and speech perception [21, 22, 20]. With clustered acquisitions an asymmetrical repetition time is employed which means that two or even more dynamic scan volumes are acquired instantly following stimulation. A subsequent silent interval then occurs which is sufficiently long to allow the hemodynamic response return to steady state. Presentation of auditory stimuli in silence is advantageous as acoustic gradient noise has been demonstrated to interfere with proper sensory and cognitive processing and thus may be considered a potentially detrimental confound. Thus, the CTA technique can be considered a beneficial fMRI acquisition protocol with considerable potential for application in auditory fMRI studies. The combination of an enduring silent period, followed by rapid acquisition of multiple brain scans, could be considered an optimal compromise between two conflicting demands. First, the requirement to deliver auditory stimuli in a silent environment and
secondly, the need to acquire multiple succeeding volumes to collect sufficient samples to adequately model the hemodynamic response [41].

We set out the present study to provide further evidence that the CTA is a suitable tool to study speech perception by systematically comparing this approach to noisy continuous acquisition schemes. We observed functional activation in perisylvian regions which were either stronger with the CA or with the CTA approach. We discuss the results and its potential implications, particularly with respect to functional lateralization in turn.

Our results show that the CTA relative to the CA approach corresponded to stronger functional responses collected from the bilateral supratemporal plane which accommodates the entire auditory cortex in the human brain. Generally, these results are consistent with previous studies having demonstrated that the adverse effects of acoustic scanner noise were most apparent in auditory fields [18, 47, 11, 48, 42, 15, 49]. This finding suggests that signal detection in primary and secondary auditory cortices is harmed at the physiological level as continuous gradient noise itself should be considered a persistent acoustic stimulus which evokes a saturation of the neuronal population residing in the supratemporal plane. Thus, noisy continuous fMRI increases the BOLD baseline activation in auditory regions which yields smaller BOLD responses to sound stimuli when compared to a CTA scheme which benefits from a silent baseline. Corroborating evidence for the view that a noisy environment detrimentally affects neuronal functions in the context of auditory processing comes from an MEG-study which investigated the influence of noise on brain responses to spoken sentences [24]. The authors of this study report that fMRI noise presented with 75 dB sound intensity disrupted early auditory processing only in the right hemisphere. Our results concur with this finding in that the present ROI analyses discovered relatively stronger functional activation in the right relative to the left mid and anterior temporal lobe. Specifically, the right anterior and the mid portions of the STP (PP, HG) exhibit stronger brain responses for the "silent" vs. the "noisy" fMRI protocol. In other words, we observe differentially lateralized signal strength varying as a function of the acquisition technique which may gain particular relevance in the realm of fMRI studies on speech perception. In the presence of background scanner noise syllables and speech stimuli are typically perceived as relatively flat sounds due to spectral overlap of typical MR gradient noise and spoken language on one hand [50] and a stapedial muscle reflex on the other hand, that might alter the perceived sound levels and spectral characteristics of the stimulus [12]. Additionally, a reduced functional response in the right
hemisphere during continuous as compared to "silent" scanning might emerge from the lateralized processing of distinct acoustic features, with temporal sound features being apparently more strongly analyzed in the left and spectral sound features more in the right hemisphere [51-57]. By contrast, the most posterior compartment of the STP, the planum temporale, displayed a general functional leftward asymmetry irrespective of acquisition technique. As apparent from recent imaging studies the left PT appears be preferentially driven by brief phonological cues which are considered the acoustic foundation of speech [58, 59, 20]. Due to its acoustic composition, the scanner noise apparently less strongly disrupts the perception of rapidly changing cues and thus it comes as no surprise that left PT functions are not harmed by continuous fMRI. Thus, the present finding may add to the view of functional segregation of STP according to the notion of primary/basic and secondary/higher auditory perception. Furthermore, this finding of a noise disturbance effect in the anterior and mid auditory regions may have some implications for future fMRI research as it cannot be ruled out that former investigations which applied CA to examine auditory and speech perceptions may have apparently underestimated the cardinal role the right hemisphere plays in this domain.

Our comparison between the two acquisition techniques furthermore revealed stronger responses to continuous scanning, namely in the bilateral insula and in the anterior cingulum which we discuss in turn. A number of recent imaging studies have observed the anterior insula and the adjacent frontal operculum as supporting auditory perception, mainly in the context of demanding top-down processing [60-65]. Furthermore it has been shown that bilateral damage to the anterior insula may result in total auditory agnosia [66], suggesting that this region is vital for auditory processing, but its specific functional role still remains elusive [67]. An explanation which is concordant with the present results has been provided by Meyer and colleagues [68]. These authors predicate their interpretation on two fMRI-studies testing the same stimulus corpus as the current study with a CA scheme and conclude that this bilateral activation "might reflect the effort to achieve a meaningful segmentation of the inflowing auditory input ..., rather than linguistic processes per se." ([68], p. 294). Here, we augment this statement with the assumption that the insular activation only observed in the CA mode indicates effortful auditory perception hampered by persistent acoustic gradient noise, rather than auditory perception per se. Since enhanced effort was necessary in the noisy scanning environment to retrieve intelligible information from the auditory stimuli, it is readily conceivable that this particular region was significantly more engaged during
continuous scanning. This interpretation is supported by the behavioural data, which showed a significantly faster response in “silent” CTA than during continuous scanning and a recent fMRI-study which demonstrates insula activity in more distractive conventional fMRI relative to a less disruptive scanning protocol [2].

Additionally, we found stronger responses in the anterior cingulum during the “noisy” vs. the “silent” fMRI. The anterior cingulum has been described as a region which supports the online monitoring of performance and unspecific aspects of cognitive control [69, 70] which indicates that a noisy environment which compromises auditory perception probably accounts for this finding as it necessitates more cognitive effort to perform the task of interest. This view also receives support by a recent study which showed that scanner noise produces activity in the anterior cingulum [71].

Taken together the present study shows that the CTA scheme can successfully be applied to map the functional response to auditory sentence stimuli of extended duration in a silent environment. Advantageously, CTA combines fMRI scanning devoid of interference with scanner acoustic noise and time-efficient data collection. At the physiological level we show that the CTA is associated with stronger functional activation in the auditory cortices as this approach is not limited by saturation of the neuronal population evoked by persistent scanner noise. At the perceptual level, the CTA scheme does not conflict with the spectral frequencies of auditory stimuli, and especially spoken language while the acoustic noise generated by the CA particularly disrupts the processing in right auditory fields and hence should be considered a masking contamination. Thirdly, frontal regions more strongly responding during CA acquisition apparently reflect unspecific effort due to foreground/background decomposition of inflowing auditory input additionally induced in a noisy environment. In addition to these advantages the CTA mode is also more comfortable for the participants as they are not exposed to persistent aversive scanner noise which makes CTA a suitable and convenient tool for future fMRI investigation of auditory and speech perception.

CONCLUSION

The present study indicates that a “silent” CTA sequence allows for statistically efficient data collection in the absence of ambient scanner noise. In particular, the CTA approach is
particularly advantageous for the investigation of basic auditory functions for two reasons. Unlike continuous scanning, CTA is not limited by the problem of saturation of neuronal populations in the auditory cortices nor is it hampered by perceptual constraints caused by masking contamination of gradient scanner noise. Finally, speech perception measured with a CTA protocol does not recruit non-auditory areas, i.e. the insula and the cingulum. It rather seems that involvement of the latter regions more likely reflects effortful processing of inflowing auditory input which is partially obscured by ambient pulsed scanner noise rather than mirroring speech perception per se.

Acknowledgments

CS and MM are supported by the Swiss National Science Foundation grant 46234101. EG and TZ are supported by the Swiss National Science Foundation grant 46234103. The authors wish to thank for the continuing support of Philips Medical systems and the financial support by ETH Zurich (Strategic Excellence Project TH 7/02-02). Furthermore we are indebted to the Max Planck Institute of Human Cognitive and Brain Sciences Leipzig for providing the stimulus corpus. We also gratefully acknowledge Betina Schnepf and Simon Baumann for motivating discussions.

REFERENCES


CONCLUSION & FUTURE PROSPECTS

In the present thesis most recent advances in MRI were transposed into beneficial instruments for neuroscientific brain research and their potential was evaluated in several studies. Generally, the optimal fMRI acquisition protocol aims for highest functional sensitivity at concomitantly finest spatial and temporal resolution, together with artefact-free image quality – and for fMRI in audition a silent scanning environment. Yet, all these features are interrelated and, as the best protocol which fulfills all these requirements at the same time obviously does not exist, compromises have to be made. In the framework of its specific requirements this thesis aimed identifying the best compromise for several fMRI studies encompassing distinct regions of the brain.

The studies described in this work endorse that fMRI has emerged a versatile, ease-of-use and widespread available technique, holding high potential for a multitude of applications. FMRI in combination with the most technical advances such as parallel imaging has become an invaluable tool for a variety of applications with some of the eminent obstacles associated with gradient-echo planar acquisition partly surmounted: The enhanced encoding speed as compared to a conventional full Fourier encoding can be used either to reduce susceptibility-related artefacts at preserved spatial resolution due to shorter acquisition time, or to enhance the spatial resolution at preserved imaging time. Furthermore, at preserved total acquisition time, it allows more time for the imagining gradients to build up and hence to reduce the acoustic scanner noise. Finally, sophisticated task and acquisition designs, such as sparse or clustered sparse sampling, have been shown promising strategies for fMRI devoid of distracting scanner acoustic scanner noise.

However, since its very beginning and despite its intensive use in cognitive brain research, neuropsychology and psychiatry, medical diagnostics and treatment planning, some pertinent methodological issues still remain in functional MRI. Due to its predominantly vascular nature the functional signal is smooth on a temporal and spatial scale. The amplitude of the (positive) BOLD response usually builds up during 4-6 s and eventually ceases during another 12-30 s. The intrinsic spatial resolution is determined by the underlying vascular capacity and architecture and the spatial resolution limit is supposed to be the 'venous unit' [1]. Furthermore, 'downstream activation', i.e. oxygenation changes draining from the site of
activation down the vascular tree and accumulating in extracerebral veins, as well as intravascular signal changes have to be carefully separated from the BOLD signal originating from the actual site of the activation. Notably, previous studies used the initial negative, rather than the positive BOLD response to localize activation with a spatial specificity in the order of visual cortical columns [2]. This early response might supposedly relate to oxygen extraction, and hence an increase in deoxyhemoglobin content, due to increased metabolic demand concomitant with neuronal activation prior to the onset of the vascular response. It therefore might be more restricted to the actual site of activation. However, its exact physiological correlates and mechanism are not yet conclusive [3, 4]. Moreover, since the BOLD signal is dependent on the absolute deoxyhemoglobin content, it is driven by cerebral blood flow, vascular volume changes as well as changes in oxygen consumption, which complicates its interpretation. In particular, the BOLD signal increases with cerebral blood flow, because more oxygenated blood is pushed through the venous bed, but it decreases with increasing blood volume and oxygen consumption alike, since the total amount of deoxyhemoglobin increases. Hence, if only a small BOLD signal is detected, it might be either correspond to only a small increase in blood flow and/or a large increase in blood volume and oxygen consumption. It therefore remains ambiguous, whether this reflects a large or small functional 'activation', except if these physiological parameters were also recorded and/or physiological relationships among these assumed [5].

Sophisticated investigation and application of functional contrasts other than blood oxygenation level dependent might hence become interesting future perspectives in functional neuroimaging. Cerebral blood flow (CBF) imaging already started entering routine praxis, and the non-invasive measurement of cerebral blood volume (CBV) can now reliably be performed with the recently developed VASO (vascular space occupancy) technique [6]. While both relate more directly to the 'pure' physiological measures, the underlying functional contrasts still provide only an indirect measure of neuronal activity and hence come with the temporal and spatial characteristic and limitation of a vascular response. Yet, activation-induced increase in extra-cellular water content, tentatively ascribed to a transient swelling of cortical cells, has recently been put forward as to serve another functional contrast, that can be measured using diffusion-weighted MRI [7] or as SEEP (signal enhancement of extravascular protons [8]). Finally, and although not a functional contrast per se, diffusion tensor

---

6 (Malonek and Grinvald, Science 1996: '...watering the whole garden for the sake of one thirsty flower.')
or q-space diffusion-weighted MR imaging (e.g. [9, 10]) allow to collect information about the course of major nerve fiber tracts and hence to bridge the gap between brain function and neuroanatomy.

REFERENCES


LIST OF PUBLICATIONS


- Prof. Dr. Peter Bösiger, der diese exzellente, motivierende wissenschaftliche Umgebung geschaffen hat und der immer an mich und meine Arbeit geglaubt und diese mit allen Mitteln unterstützt hat. Es ist sein Verdienst, dass in unserer Gruppe ein so frisches, menschliches und freundliches Klima herrscht
- Prof. Dr. Lutz Jäncke für das Koreferat und Martin Meyer, der mir immer geduldig zur Seite stand, wenn ich in meiner Arbeit steckenblieb oder an Lösungsansätzen zweifelte. Mit wissenschaftlicher Expertise, organisatorischem Geschick und immer um faires, hilfsbereites Miteinander bemüht, hat er unsere kleine Auditorikgruppe geformt
- Klaas, der die Welt der MR-Physik auf ebenso faszinierende wie klare Weise darzustellen weiss und uns immer neue Wunderwelten zu erschliessen vermag; sowie Jeff, dessen blitzschneller Geist uns immer wieder verblüffen liess. Beiden verdanke ich Geleit und Halt auf meinen ersten wackeligen Schritten in der parallelen Bildgebung
- PD Dr. Katrin Henke, die mich anzuspornen und zu begeistern vermochte durch ihren mitreissenden Optimismus, ihr Interesse für technische Neuerungen, die sie direkt in ihre Arbeit einfließen liess
- Prof. Dr. Alumit Ishai for guidance, supervision and steady support. She opened my eyes for the beauty of our brains – and faces, and she introduced me to the foundations of modern neuroscience
- Diane-Elise und Pascal, auf die die Wogen meiner Enttäuschungen und Rückschläge meist unmittelbar niedergingen, und mit denen ich nächtelang Dinnieren, Diskutieren und DVDs schauen konnte
- Roger, unser aller Schutzpatron, der mit ausgefeiltesten Tricks und unglaublicher Zähigkeit alle Computerprobleme zu lösen vermag und als guter Geist des MR Zentrums immer da war, wenn ich hilflos in der Klemme sass
• meinen Büro-Gspönli für die vielen heiteren, arbeitsamen und motivierenden gemeinsamen Stunden: Diane-Elise, meinem treuen Gegenüber, Carolin, meiner Leidensgenossin beim SPM-Auswerten, Anke, deren Energie und Überschwang ansteckend sind, Marco und Marco, die oft ungewollt für filmreife Situationskomik sorgten, Thomas L. und Berti, die Ausgeglichenheit verströmten, Betina, die mich immer gern zur Kaffeemaschine begleitete und immer bereit war, an fachlichen Fragestellungen mitzuknobeln und mit der ich herrlich diskutieren konnte
• Michi S., Christof und Philipp, die immer ein offenes Ohr für mich hatten
• Ulrike und Gerard, die mich mit meiner Diplomarbeit in die Gruppe und in das wissenschaftliche Arbeiten eingeführt haben
• Salome, Mike und Dieter, die Sonne und Leben ins MR Zentrum brachten und dafür sorgten, dass immer alles funktioniert
• Rolf, Basti, Nicola, Thomas J., Hendrik, Andrea, Stigi, Reto, Christoph, Florian, Urs, Martin, Oliver, Michi H. und Moe, die meine Zeit am IBT begleitet haben, sowie Urs Sturzenegger, der immer eine schützende Hand über unsere Scanner hielt
• Kai, Simon, Evelin, Tino, Elke und Susan aus der "Jäncke-Gruppe", die mich mit offenen Armen aufnahmen, auch wenn ich dem IBT fest verwurzelt blieb, für viele anregende Diskussionen und die tolle Zusammenarbeit an den gemeinsamen Projekten
• meinen Freunden, insbesondere Corina, die in unerschütterlicher Weise sowohl die Geheimnisse des Weltalls als auch des Lebens zu ergründen sucht und mich zum ruhigen Nachdenken in meinem oft hektischen Alltag bewegt hat
• meiner WG, die mir alle Freiheiten gewährte und dem 'Uhu' ein ausgesprochen komfortables und bequemes Nest bot
• Ingo, der mich in seiner unverkennbaren Waschkies-Art immer wieder zum Lachen bringt und riesengrosse Probleme ganz winzig klein machen kann
• meinen Eltern, die mich immer und in jeder Hinsicht unterstützt, und mit einer schier unendlichen Geduld all meine Macken ertragen haben
CURRICULUM VITAE

I was born on July 30th 1976, in Leipzig, Germany, as the daughter of Irene and Eckard Schmidt. I grew up in Leipzig-Marienbrunn, where I attended primary school. In summer 1995 I graduated with the Abitur from the Koenig-Albert Gymnasium Leipzig.

In autumn 1995 I started my studies in Physics at the University of Leipzig. In 1997 I went to the Swiss Federal Institute of Technology (ETH) Zurich, Switzerland, as an exchange student of the European Mobility Scheme for Physics Students programme. During this year I specialized in Biophysics and Medical Physics and had the opportunity to work under the supervision of Prof. Dr. K. Wuethrich, Institute for Molecular Biology and Biophysics, on "Quantitative Measurement of $^{3}$JN.Hß coupling constants in $^{15}$N labeled BPTI: Reparametrisation of the Karplus Equation" and under Prof. Dr. P. Ruegsegger, Institute for Biomedical Engineering, on "Microtomography in teeth". At the end of my exchange year at the at the Swiss Federal Institute of Technology (ETH) Zurich, Switzerland, I passed the diploma exams with specializations in Biomechanics, Biophysics, and Medical Physics. From 1999-2000 I worked on my physics diploma thesis entitled 'Arterial Spin Labelling Techniques for Event-Related fMRI'. This work was supervised by PD Dr. D.G. Norris from the Max-Planck-Institute for Cognitive Neuroscience, Leipzig, Germany, and by Prof. Dr. P. Boesiger, head of the Institute for Biomedical Engineering of the Swiss Federal Institute of Technology (ETH) and University Zurich, Switzerland, where the work was performed. I graduated in October 2000 with the Diploma in Physics at the University of Leipzig.

In February 2001, I joined the group of Prof. Dr. Peter Boesiger as a PhD student and research assistant. My goal was to evaluate and optimize sequence protocols for dedicated fMRI studies and I particularly focussed on susceptibility artefact reduction. Since October 2003 I'm also associated to the group of Prof. Dr. L. Jancke, Departement of Neuropsychology, University Zurich, Switzerland, where I develop and optimize methods for the investigation of 'Short-term and long-term plasticity in the auditory system' (Swiss National Science Foundation Project 46234101).