2-2014

On the role of neuronal oscillations in auditory cortical processing

Monica Noelle O'Connell

Graduate Center, City University of New York

How does access to this work benefit you? Let us know!

Follow this and additional works at: https://academicworks.cuny.edu/gc_etds

Part of the Cognitive Psychology Commons, and the Neuroscience and Neurobiology Commons

Recommended Citation

O'Connell, Monica Noelle, "On the role of neuronal oscillations in auditory cortical processing" (2014). CUNY Academic Works.
https://academicworks.cuny.edu/gc_etds/85

This Dissertation is brought to you by CUNY Academic Works. It has been accepted for inclusion in All Dissertations, Theses, and Capstone Projects by an authorized administrator of CUNY Academic Works. For more information, please contact deposit@gc.cuny.edu.
ON THE ROLE OF NEURONAL OSCILLATIONS IN AUDITORY CORTICAL PROCESSING

by

MONICA NOELLE O'CONNELL

A dissertation submitted to the Graduate Faculty in Psychology in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York 2014
This manuscript has been read and accepted for the Graduate Faculty in Psychology in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

Charles E. Schroeder

Date

Chair of Examining Committee

Maureen O'Connor

Date

Executive Officer

Sophie Molholm, Ph.D.

John F. Smiley, Ph.D.

Donald A. Wilson, Ph.D.

A.M. Clare Kelly, Ph.D.

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK
Abstract

ON THE ROLE OF NEURONAL OSCILLATIONS IN AUDITORY CORTICAL PROCESSING

Monica Noelle O’Connell

Adviser: Charles E. Schroeder, Ph.D.

Although it has been over 100 years since William James stated that “everyone knows what attention is”, its underlying neural mechanisms are still being debated today. The goal of this research was to describe the physiological mechanisms of auditory attention using direct electrophysiological recordings in macaque primary auditory cortex (A1). A major focus of my research was on the role ongoing neuronal oscillations play in attentional modulation of auditory responses in A1.

For all studies, laminar profiles of synaptic activity, (indexed by current source density analysis) and concomitant firing patterns in local neurons (multiunit activity) were acquired simultaneously via linear array multielectrodes positioned in A1. The initial study of this dissertation examined the contribution of ongoing oscillatory activity to excitatory and inhibitory responses in A1 in passive (no task) conditions. Next, the function of ongoing oscillations in modulating the frequency tuning of A1 during an intermodal selective attention oddball task was investigated. The last study was aimed at establishing whether there is a hemispheric asymmetry in the way neuronal oscillations are utilized by attention, corresponding to that noted in humans.

The results of the first study indicate that in passive conditions, ongoing oscillations reset by stimulus related inputs modulate both excitatory and inhibitory components of local neuronal ensemble responses in A1. The second set of experiments demonstrates that this mechanism is utilized by attention to modulate and sharpen frequency tuning. Finally, we show that as in humans, there appears to be a specialization of left A1 for temporal processing, as signified by greater temporal precision of neuronal oscillatory alignment. Taken together these results underline the importance of neuronal oscillations in perceptual processes, and the validity of the macaque monkey as a model of human auditory processing.
Dedication

This dissertation is dedicated to my family; Peter, Lili & Nora, whose unconditional love, support and reassurance have made all the difference during this long and eventful academic journey. I am truly lucky and forever grateful.
Acknowledgements

First of all, I would like to express my extreme gratitude to my mentor Dr. Charles Schroeder, not only for his outstanding support, guidance and encouragement on this project, but also for his warm and invaluable friendship throughout the years. I also thank Dr. Peter Lakatos for his unlimited willingness to advise on and steer the analysis of this data, as well as his immense assistance in writing Matlab scripts.

Last but not least, many thanks to the fabulous Schroeder lab members at NKI: Tammy McGinnis, Annie Barczak, Arnaud Falchier, Yoshi Kajikawa and Debbie Ross, who have all made meaningful contributions to my graduate studies and with whom I've had "mighty craic". I love you all and wish nothing but the best for your futures.
# Table of Contents

**GENERAL INTRODUCTION** .................................................................................................. 1

1.1. PHASE RESET VS. EVOKED TYPE RESPONSES IN SENSORY CORTICAL AREAS .............. 4
1.2. THE EFFECT OF STIMULUS FEATURES ON EVENT RELATED RESPONSES ...................... 8
1.3. THE MECHANISM OF AUDITORY ATTENTION IN PRIMARY AUDITORY CORTEX .......... 9
1.4. HEMISPHERIC SPECIALIZATION OF AUDITORY FUNCTION ........................................... 11
1.5. OBJECTIVES ................................................................................................................... 15

**CHAPTER 1 - DUAL MECHANISM OF NEURONAL ENSEMBLE INHIBITION IN PRIMARY AUDITORY CORTEX** ........................................................................................ 17

2.1. SUMMARY ...................................................................................................................... 18
2.2. INTRODUCTION .............................................................................................................. 19
2.3. RESULTS ......................................................................................................................... 22
   2.3.1. Properties of MUA Suppression in Inhibitory Sidebands ........................................ 23
   2.3.2. Laminar Profiles of BF vs. non-BF Inhibitory Responses in A1 ................................. 27
   2.3.3. The Physiological Mechanism of the Inhibitory CSD Response .............................. 29
2.4. DISCUSSION ................................................................................................................... 37
   2.4.1. Prevalence of Inhibition in Primary Auditory Cortex ............................................. 37
   2.4.2. Mechanisms of Inhibitory Responses .................................................................... 39
   2.4.3. Conclusions ............................................................................................................ 45
2.5. EXPERIMENTAL PROCEDURES .................................................................................. 45
   2.5.1. Subjects .................................................................................................................. 45
   2.5.2. Surgery .................................................................................................................. 46
   2.5.3. Electrophysiology .................................................................................................. 46
   2.5.4. Data analysis ......................................................................................................... 48

**CHAPTER 2 - SHARPENING OF FREQUENCY TUNING BY SELECTIVE ATTENTION IN PRIMARY AUDITORY CORTEX** ................................................................. 52

3.1. SUMMARY ...................................................................................................................... 53
3.2. INTRODUCTION .............................................................................................................. 54
3.3. RESULTS ......................................................................................................................... 57
   3.3.1. Sharpening of Frequency Tuning By Attention ....................................................... 58
   3.3.2. Entrainment of Delta Oscillations by Attended Tone Streams ............................... 63
   3.3.3. The Net Effect of Attention on the Excitability of A1 Neuronal Ensembles ............... 67
3.4. DISCUSSION ................................................................................................................... 70
   3.4.1. Mechanism of Increased Response Selectivity in A1 ............................................ 71
   3.4.2. Layer Specificity of Attention Effects ................................................................. 73
3.4.3. The Effect of Attention on Auditory Stimulus Processing ................................................. 75
3.4.4. Context and Content ................................................................................................. 78
3.4.5. Conclusions ............................................................................................................ 79
3.5. EXPERIMENTAL PROCEDURES ....................................................................................... 79
3.5.1. Subjects .................................................................................................................. 79
3.5.2. Surgery .................................................................................................................. 80
3.5.3. Electrophysiology ................................................................................................. 80
3.5.4. Behavioral task and stimuli ..................................................................................... 81
3.5.5. Data analysis ......................................................................................................... 83

CHAPTER 3  -  MULTI-SCALE ENTRAINMENT OF COUPLED NEURONAL
OSCILLATIONS IN PRIMARY AUDITORY CORTEX ................................................... 86

4.1. SUMMARY ...................................................................................................................... 87
4.2. INTRODUCTION .............................................................................................................. 88
4.3. RESULTS ......................................................................................................................... 90
4.3.1. The Effect of Engagement on Responses to Click Trains ............................................ 92
4.3.2. The Pattern of Delta and Gamma Frequency Entrainment across A1 ......................... 96
4.3.3. Engagement Related Hemispheric Asymmetry ........................................................... 103
4.4. DISCUSSION ................................................................................................................. 105
4.4.1. Mechanisms of Predictive Response Suppression ......................................................... 106
4.4.2. The Effect of Engagement on Neuronal Activity in A1 ............................................... 107
4.4.3. The Importance of Broadband Transient Sounds in Auditory Processing .................... 108
4.4.4. Evidence for Hemispheric Functional Lateralization in Non-Human Primates .............. 111
4.4.5. Conclusions .......................................................................................................... 113
4.5. EXPERIMENTAL PROCEDURES ..................................................................................... 113
4.5.1. Subjects ................................................................................................................ 113
4.5.2. Surgery ................................................................................................................ 114
4.5.3. Electrophysiology ................................................................................................... 114
4.5.4. Behavioral task and stimuli ...................................................................................... 115
4.5.5. Data analysis ......................................................................................................... 117

GENERAL DISCUSSION ..................................................................................................... 120

5.1. RELATIONSHIP OF STUDIES ......................................................................................... 123
5.2. HEMISPHERIC LATERALIZATION OF AUDITORY FUNCTION ........................................ 125
5.3. NEURONAL OSCILLATIONS AND BRAIN DISORDERS ................................................... 126
5.4. FUTURE DIRECTIONS ................................................................................................... 129

REFERENCES ....................................................................................................................... 132
# List of Figures

Figure 1. *Excitatory and inhibitory MUA responses (Chapter 1)* .............................................. 23

Figure 2. *The frequency dependence of granular MUA response onset (Chapter 1)* .............. 25

Figure 3. *MUA and current source density (CSD) profiles associated with BF and non-BF tones in A1 (Chapter 1)* .......................................................... 27

Figure 4. *Laminar CSD response onset in excitatory (BF) and inhibitory response profiles (Chapter 1)* .............................................................................................. 29

Figure 5. *CSD amplitude profiles (Chapter 1)* ............................................................................ 30

Figure 6. *Oscillatory properties of BF and non-BF responses (Chapter 1)* ............................ 32

Figure 7. *Delta oscillatory entrainment (Chapter 1)* ................................................................. 35

Figure 8. *Phase triggered averages of spontaneous CSD and concomitant recorded MUA (Chapter 1)* ......................................................................................... 43

Figure 9. *Representative laminar CSD and MUA profiles in response to attended and ignored BF and non-BF streams in A1 (Chapter 2)* .............................................. 59

Figure 10. *The effect of attention on the MUA response and sharpness of tuning (Chapter 2)* .................................................................................................................. 60

Figure 11. *Delta entrainment to different frequency tone streams (Chapter 2)* ...................... 64

Figure 12. *Pooled delta phase measures (Chapter 2)* ............................................................... 66

Figure 13. *Modulation of pre-stimulus excitability (Chapter 2)* ............................................ 69

Figure 14. *The auditory stimulus stream (Chapter 3)* ............................................................ 91

Figure 15. *Engagement in a task generally suppresses responses to click trains (Chapter 3)* ......................................................................................................................... 93

Figure 16. *Representative laminar response profiles (Chapter 3)* ........................................ 95

Figure 17. *The amplitude of spontaneous and auditory stimulus stream related delta and gamma band neuronal activity (Chapter 3)* ....................................................... 97

Figure 18. *The distribution of pre-stimulus delta and gamma oscillatory phases across differently tuned A1 regions (Chapter 3)* ................................................................. 99

Figure 19. *MUA responses in A1 regions with differing pre-stimulus phases (Chapter 3)* .......................................................................................................................... 102

Figure 20. *Hemispheric lateralization of intertrial phase coherence and the amplitude of*
delta and gamma oscillations

(Chapter 3)
GENERAL INTRODUCTION

Ever since the identification and description of the different brain rhythms (e.g. alpha 8-13 Hz) in the electroencephalogram (EEG) by Hans Berger in the early 20th century (Berger, 1929), it has been proposed that spontaneous brain waves or oscillations of the brain reflect its internal “state” and may influence its response to incoming sensory stimuli, as neuronal oscillations signal rhythmic variability of neural excitability along temporal and spatial dimensions (Buzsaki and Draguhn, 2004; Fries, 2005; Lakatos et al., 2005b). As early as 1933 George Bishop noted the physiological importance of the brain's ongoing oscillatory activity, as while stimulating an optic nerve he detected recurring or cyclical excitability variations in the visual cortex of the rabbit (Bishop, 1933). Thus he wrote “In general, it is not necessary to infer that each individual impulse traveling up a fiber from the retina arrives as a unit impulse in the cortex, and registers there as such. Rather, we would look upon the cortex as being in constant activity, the physiological activity of the whole network of neurons bearing some direct relationship to the ‘present state’ of the animal’s complex behavior which is sometimes referred to as his ‘mental state”.

Following in these early pioneers footsteps, mid-century researchers first discovered that reaction times varied depending on the phase of intrinsic neuronal oscillations (Lansing, 1957; Callaway, III and Yeager, 1960; Callaway, III and Alexander, Jr., 1960), and differences in amplitudes and latencies of event related response components related to the phase of oscillatory activity at which the stimuli were presented were also reported in the subsequent years (Callaway, III and Yeager, 1960; Bechtereva and Zontov, 1962; Dustman and Beck,
1965; Remond and Lesevre, 1967). Nonetheless, until recently the notion that internal rhythmic excitability fluctuations merely represent “noise” in the brain still prevailed. However, in the last couple of decades increasingly more studies were able to demonstrate mechanistic roles for neuronal oscillations in various brain structures (Brandt et al., 1991; Arieli et al., 1996; Polich, 1997; Fries et al., 2001a; Fries et al., 2001b; Lakatos et al., 2005b; Womelsdorf et al., 2006b; Jensen et al., 2007; Lakatos et al., 2008; Busch et al., 2009; Lakatos et al., 2009; Besle et al., 2011; Dugue et al., 2011), thereby making it more difficult to deny that oscillations are random excitability fluctuations, but rather they reflect the organized internal electrophysiological context of the brain, that is orchestrated for the most efficient processing of external and internal information (Buzsaki and Chrobak, 1995; Lakatos et al., 2009). For this reason, the notion that ongoing neuronal oscillations have a direct effect on sensory processing and thus perception, and are not merely background "noise" slowly took hold.

A key feature of neuronal oscillations is that they have both high and low excitability phases, that rhythmically alternate, and during which the oscillating neuronal ensemble’s response to a sensory input is either amplified (during high excitability phase) or suppressed (during low excitability phase) (for a review see Schroeder et al., 2008). If these phases occurred independent of the timing of inputs, they would randomly enhance or suppress responses, resulting in an unstable sensory representations. In this case, oscillations could be regarded as noise from the perspective of sensory information processing. However, even as early as 1974 (Sayers et al., 1974) the mechanism of oscillatory phase reset by sensory inputs was proposed by which the brain aligns a specific phase of its neuronal oscillations to sensory inputs. This view did not genuinely take hold until the controversial paper by Makeig (Makeig et al., 2002), in which he states that sensory ERPs are attributable to the reorganization or reset of the phases of oscillations in particular frequency bands by sensory related inputs. This
sparked a decade long debate, in which research groups set out to determine what – if any – portion of the sensory ERP is due to phase reset (Brandt, 1997; Basar et al., 1999; Makeig et al., 2002; Penny et al., 2002; Jansen et al., 2003; Kruglikov and Schiff, 2003; Rizzuto et al., 2003; Yamagishi et al., 2003; Fell et al., 2004; Klimesch et al., 2004; Shah et al., 2004; David et al., 2005; Gruber et al., 2005; Hamada, 2005; Makinen et al., 2005; Mazaheri and Picton, 2005; Fuentemilla et al., 2006; Mazaheri and Jensen, 2006; Naruse et al., 2006; Hanslmayr et al., 2007; Klimesch et al., 2007; Sauseng et al., 2007; Barry, 2009). The pure phase resetting model specifically maintains that there is simply a phase resetting of ongoing oscillations to a specific value in each trial, without any amplitude increase in the post stimulus timeframe compared to the baseline. As a result post stimulus EEG oscillations are aligned or “phase-locked” across trials, thus positive and negative peaks do not average out but are detectable in the averaged responses as peaks and troughs of the ERP (Sayers et al., 1974; Basar et al., 1980; Makeig et al., 2002; Klimesch et al., 2007). This theory is in conflict with the long held assumption that responses are newly generated in response to the stimulus (thus the phrase “evoked response”), are overlaid on the ongoing EEG and are characterized by an amplitude increase from the pre to post stimulus timeframe in each trial (Dawson, 1950; Jervis et al., 1983; Makinen et al., 2005). In recent years, the debate became less polarized in light of intracortical studies demonstrating that both phase reset and evoked type neuronal activity contribute to the ERP, and there is a nonlinear relationship between the two, in that reset neuronal oscillations can modulate the evoked part of the response (Makinen et al., 2005; Lakatos et al., 2005b; Lakatos et al., 2007; Lakatos et al., 2008; Lakatos et al., 2009; Lakatos et al., 2013a). There is also recent evidence that ERPs recorded on the scalp are a mixture of evoked activity and phase reset of ongoing oscillatory activity (Makeig et al., 2004; Fuentemilla et al., 2006; Min et al., 2007; Barry, 2009; Telenczuk et al., 2010; Gomez-Ramirez et al., 2011), although the ratio is still
debated, since due to the volume conduction and summation of synchronous neuronal activity, even pure phase reset involving several different neuronal ensembles can present as an evoked type, “added” post-stimulus activity (Sauseng et al., 2007). Nonetheless, due to the elimination of volume conduction and because of high spatial resolution, intracortical recordings on the mesoscopic scale can distinguish between stimulus related phase reset and evoked type activity. Below we will describe the properties that help differentiate between the two different response types and their differing function.

1.1. PHASE RESET VS. EVOKED TYPE RESPONSES IN SENSORY CORTICAL AREAS

One of the first studies to show pure phase resetting, was an investigation of multisensory interactions in primary auditory cortex (A1) of awake macaque monkeys using somatosensory and auditory stimuli (Lakatos et al., 2007). This study demonstrated that somatosensory related responses in A1 are characterized by a low amplitude supragranularly weighted current source density (CSD) modulation combined with no transient multiunit activity (MUA) correlate, indicating that while somatosensory input does alter the net local neuronal excitability (signified by an organized poststimulus CSD pattern in the averaged response) it does not change the amount of net post-stimulus transmembrane current, and thus does not “trigger” neuronal firing. As opposed to this, the same study, like many previous ones (e.g. (Schroeder et al., 1998; Schroeder et al., 2003), demonstrated that auditory, preferred modality stimuli result in an increase of net transmembrane current in all cortical layers, amounting to an evoked type response (see above). This type of response is also characterized by a phasic MUA response due to increased synaptic currents. Thus phase-reset and evoked type responses are functionally different: while the former is modulatory, since it does not trigger suprathreshold neuronal
activity (firing), the latter is a “driving” type, which results in significant increases in the amplitude of net transmembrane currents (indexed by CSD increase) and consequently a significant increase in post-stimulus MUA indicating that specific information (e.g. frequency, location, intensity, etc.) about the stimulus is being transmitted. Lakatos and colleagues (Lakatos et al., 2007; Lakatos et al., 2009) found that these two functionally different types of responses also have different signatures in the time-frequency domain: phase reset is restricted to the phase modulation of dominant, delta, theta and gamma band oscillatory activity resulting in a biased post-stimulus phase distribution (“phase-locking”), accompanied by little or no pre-to poststimulus increase in oscillatory amplitude, which is the hallmark of oscillatory phase resetting (Makeig et al., 2004; Shah et al., 2004). As opposed to this, evoked type or driving responses were characterized by a sharp onset de-novo generated waveform, which in the time frequency domain is represented by a spectrally distributed amplitude increase, coupled with biased phase distribution similarly spanning the whole spectrum due to the sharp onset (similar to the spectrum of a delta function or square wave). It is important to note that this does not mean that auditory, preferred modality stimuli do not reset ongoing oscillations, but the signatures of phase reset are “masked” by the larger amplitude evoked activity. A follow up study by Lakatos and colleagues (2009), illustrated that like somatosensory, visual stimulus related inputs are also capable of modulating A1 neuronal activity via the phase resetting of ongoing oscillations (Lakatos et al., 2009). Importantly, this study additionally showed that phase reset of ongoing oscillatory activity only accompanied visual and auditory inputs in conditions when the subjects were attending to these stimuli. Thus, the authors concluded that in contrast to evoked type responses, phase reset modulation requires stimuli that are either attended or inherently salient.
By now a wealth of animal and human studies have shown that neuronal excitability is inexorably linked with oscillatory phase (Monto et al., 2008; Lakatos et al., 2008; Busch et al., 2009; Mathewson et al., 2009; Whittingstall and Logothetis, 2009). It follows that the phase to which sensory related inputs reset ongoing oscillations in A1 to, determines the effect of phase reset on subsequent, driving type responses: if the ongoing activity of neuronal ensembles is reset to a high excitability phase (i.e. neurons are more excitable or closer to their firing threshold) driving responses get amplified. If however oscillations are reset to their low excitability phase, neurons are relatively hyperpolarized which results in response attenuation. Lakatos et al. (2007) found that somatosensory stimulation contralateral to the recording site reset ongoing oscillations in A1 to their high excitability phase, which when delivered simultaneously with auditory stimuli produced an enhanced auditory response. In contrast ipsilateral somatosensory stimuli reset ongoing oscillations to their opposite or low excitability phase and when paired with auditory stimuli resulted in a suppressed auditory responses. In addition to the "reset phase", a key factor determining the effect of phase reset on a subsequent driving type input is the temporal relationship between phase reset and input-evoked response. For example, there are “optimal times” following the resetting event (i.e. somatosensory input) for multisensory enhancement, which in the case of contralateral stimuli (that reset oscillations to their high excitability phase), correspond to the immediate post-reset and then to periods of delta-, theta- and gamma-band EEG oscillations. The combined results of these studies confirm that the phase of ongoing oscillations modulates sensory information processing, and that in order to optimally process the sensory environment, oscillatory phases are orchestrated by a reset mechanism that is a slave to the goals of the observer. These studies also confirm a previous suggestion by Buzsaki and Chrobak (Buzsaki and Chrobak, 1995) that ambient rhythmic activity composed of “subthreshold" neuronal oscillations forms the
electrophysiological context that allows and directs the processing of sensory content that is encoded by suprathreshold neuronal activity (spikes).

Consistent with these results, a recent human behavioral study suggested that cross-modal phase reset of ongoing neural oscillations in visual cortices by an auditory stimulus could be the mechanism that enhances visual-target detection (Fiebelkorn et al., 2011). Near-threshold visual stimuli were presented following an auditory stimulus at intervals varying from 0ms (simultaneous with sound) up to 6000 ms. The authors found that performance fluctuated in a manner that was time-locked to the presentation of the sound, at periods that corresponded to the wavelength of delta oscillatory activity. Since EEG was not recorded during the task, the authors could not determine within which oscillatory frequency bands phase resetting was occurring. Nevertheless the delta periodicity is suggestive of a dominance of low frequency neuronal oscillations on perceptual outcomes. Another human study has shown that like delta, the theta-alpha band oscillatory activity can also influence behavioral performance (Thorne et al., 2011). This study provided EEG evidence for increased theta-alpha phase locking in auditory cortex to pure tones only when preceded 30-75 ms by visual stimuli. Similarly, Romei and colleagues demonstrated cross-modal phase resetting of alpha band (~10Hz) oscillations in visual cortex by sounds, which appeared to improve detection of phosphenes induced by occipital transcranial magnetic stimulation (TMS) in a cyclical (10Hz) manner (Romei et al., 2012). Overall these results highlight the important effect of modulatory inputs on cortical excitability and thus on the processing of subsequent driving inputs, and in consequence - most importantly- on perceptual outcomes and behavioral performance.
1.2. THE EFFECT OF STIMULUS FEATURES ON EVENT RELATED RESPONSES

Lakatos et al., (2007; 2009) showed that in primary cortices, heteromodal (non-preferred modality stimuli) produce solely modulatory responses via the phase resetting of ongoing oscillations within dominant frequency bands. This detail prompted us to examine the types of activation of A1 by preferred modality but non-optimal stimuli e.g. pure tone whose frequency does not match the preferred tuning of a cortical column, since the most fundamental organizing dimension of A1 neuronal ensembles is stimulus frequency. A1 has an orderly and progressive spatial arrangement of tone frequency neural representations (tonotopic map); as one progresses in an anterolateral to posteromedial direction frequencies systematically increase from low to high (Kaas and Hackett, 2000;Hackett, 2011), which is attributed to direct, spatially organized thalamocortical (TC) inputs from the ventral subdivision of the medial geniculate nucleus (MGNv) of the thalamus (Huang and Winer, 2000;Lee et al., 2004;Liu et al., 2007). If the frequency of a given auditory stimulus matches the frequency tuning of a neuronal ensemble, this “best frequency” (BF) stimulus results in a suprathreshold activation of the neuronal ensemble or an evoked response, accompanied by phase reset if stimuli are not ignored. Non-BF tones, which differ in frequency from the tuning of a neuronal ensemble will not result in suprathreshold activation due to the lack of specific, driving thalamocortical inputs activating glutamatergic layer 4 neurons. In fact, in some cases, these tones can result in an inhibition of firing, termed sideband inhibition (Shamma and Symmes, 1985;Suga, 1995;Sutter et al., 1999). The question that inspired our first study (Chapter 1) was: do non-BF tones result in a modulatory response and if so, is it in anyway different from the modulation related to BF tones. Our linear array multielectrodes allow us to sample laminar LFP and MUA profiles directly from A1. CSD analysis of the LFP profile indexes the location, direction, and density of transmembrane current flow (which is the first-order neuronal response to synaptic input),
eliminates the effects of volume conduction and provides a sensitive measure of synaptic activity even in cases of subthreshold, modulatory responses (Nicholson, 1973; Schroeder et al., 1998). Thus, we can examine both the ongoing and the stimulus related activity of neuronal ensembles independently of neighboring activity.

Since inhibitory responses in A1 are not characterized by an MUA increase but rather a decrease we speculated that the modulation of ongoing cortical oscillations via their phase resetting to a low excitability phase may, in theory, assist in the feed-forward inhibition mediated by TC afferents from MGNv (Wehr and Zador, 2003; Zhang et al., 2003; Tan et al., 2004; Wehr and Zador, 2005; Wu et al., 2008). Thus in Chapter 1 we 1) investigate whether inhibitory responses (distinguished by the largest stimulus related MUA suppression) to pure tones show signatures of modulatory (i.e. phase reset) or evoked type responses, 2) determine if auditory stimulus frequency affects the phase oscillations are reset to in any given frequency selective A1 neuronal ensemble.

1.3. THE MECHANISM OF AUDITORY ATTENTION IN PRIMARY AUDITORY CORTEX

For decades it has been known that attention toward a specific feature of a stimulus (e.g. location or frequency) enhances that feature's neural representation compared to when that same stimulus is ignored. The majority of these findings resulted from research done on the mechanisms of attention in the visual modality (Harter et al., 1982; Hillyard, 1984; Spitzer et al., 1988; Corbetta et al., 1990; Heinze et al., 1994; O'Leary et al., 1996; Luck et al., 1997; Lakatos et al., 2008), complemented by similar findings in the auditory system (Hillyard et al., 1973; Alho, 1992; Woldorff et al., 1993; Alcaini et al., 1994; Grady et al., 1997; Fujiwara et al., 1998; Jancke et al., 1999; Petkov et al., 2004; Da Costa, 2013). More recent studies have shown that in
addition to response gain effects, attention can enhance response selectivity or response contrast of visual neuronal populations coding for spatial location (Womelsdorf et al., 2006b; Fischer and Whitney, 2009). Likewise in the auditory modality it has been shown that focusing attention to a specific stimulus feature (e.g. frequency) augments neuronal selectivity for that feature in the cortical region that preferentially processes it (Kauramaki et al., 2007; Okamoto et al., 2007; Ahveninen et al., 2011; Neelon et al., 2011). Kauramaki and colleagues (2007), for example, using a notch filtered masking paradigm where the width of the notch in the white noise masker was varied around the frequency of the standard tone, found an amplitude increase in the N100 ERP component in the attend condition but as a non-multiplicative function of the width of the notch when compared to the ignore condition, suggesting that during attention the increased selectivity of the neural population processing the attended tone is offsetting the suppressive effect of the masking noise. Therefore it appears that both the auditory and visual cortices may employ the same neural mechanisms during attention, specifically increasing the contrast of attended stimulus features utilizing both enhancement of relevant and suppression of irrelevant sensory information in cortical – and perhaps thalamic – neuronal ensembles.

These attentional effects are undoubtedly responsible for the task-related modulation of receptive fields of A1 neurons that have been observed in intracortical recordings in behaving animals (Edeline and Weinberger, 1993; Ohl and Scheich, 1996; Fritz et al., 2003; Fritz et al., 2005b; Fritz et al., 2007b; Atiani et al., 2009; Galindo-Leon et al., 2009). While the effects of selective attention on auditory responses have been described by numerous studies, we lack information on the physiological mechanisms by which response enhancement and suppression are achieved, thus in Chapter 2 we examine the effects of selective attention on the frequency tuning of A1 neuronal ensembles and investigate the mechanisms by which frequency tuning is
modulated in a context in which neuronal oscillations play a central role. The reason for this is twofold: first, neuronal oscillations are capable of both enhancing and suppressing the excitability of A1 neuronal ensembles (Lakatos et al., 2007; Lakatos et al., 2013a), and second, phase reset is under strong attentional control (Lakatos et al., 2009; Lakatos et al., 2013a). To examine attention related changes in tuning and their mechanism across A1 neuronal ensembles, ideally we should record the activity of differently tuned A1 neuronal ensembles simultaneously. Since our current recording technique only allows the simultaneous sampling of two cortical locations, we opted to test for changes in tuning by recording the activity of one cortical location in each hemisphere simultaneously (to test for any hemispheric differences in attention effects, see below). We presented a series of pure tones in different blocks, whose frequency roughly covers the hearing range of the macaque monkey (0.3 kHz – 32 kHz), in attended vs. ignored conditions. This allowed us to estimate attention related changes in the frequency tuning of a given neuronal ensemble. Our specific hypothesis was that attention would result in a sharpening of tuning, and that the phase (high or low excitability phase) to which ongoing oscillatory activity is reset to by attended tones (i.e. frequency-specific phase reset) would be the mechanism responsible for the majority of this effect.

1.4. HEMISPHERIC SPECIALIZATION OF AUDITORY FUNCTION

Numerous human EEG (Liegeois-Chauvel et al., 1999; Liegeois-Chauvel et al., 2001), neuroimaging (Belin et al., 1998; Zatorre and Belin, 2001; Jancke et al., 2002; Zaehle et al., 2004; Brechmann and Scheich, 2005; Jamison et al., 2006; Hyde et al., 2008; Okamoto et al., 2009a) and behavioral (Robin et al., 1990) studies have indicated that the auditory cortices in the two hemispheres are specialized such that right auditory cortical regions have a relatively finer resolution in the frequency domain while the left cortical areas exhibit a higher temporal
resolution, which is undoubtedly relevant for speech perception. One of the more influential studies on lateralized specialization and perhaps most persuasive in demonstrating differential hemispheric activation of auditory cortex comes from a positron emission tomography (PET) study by Zatorre and Belin (2001). In this study, two stimulus parameters, rate of temporal change and spectral distribution of elements within a pattern, were varied independently in the two conditions of the experiment. During the temporal condition, where the frequency separation between two tones was held constant but the duration of each tone was randomly changed, the authors found greater left core auditory cortex activation. Whereas during the spectral condition, where the temporal pattern of presentation was held constant but the number of tones sampled fluctuated, responses were lateralized to right non-primary auditory regions. Consequently from these findings a theory was suggested by Zatorre and colleagues (Zatorre et al., 2002) that there is a trade-off among hemispheres between spectral and temporal processing precision based on the "Acoustic uncertainty principle", which maintains that it is impossible to make a precise simultaneous measurement in both the frequency and time domain of a sound i.e. employing a long time window for sound signal analysis (e.g. Fourier transform) produces high frequency but low temporal resolution, while use of a short time window results in excellent temporal but poor frequency resolution (Heisenberg, 1927;Joos, 1948).

Similarly, based on the same principle, Poeppel suggested that auditory information is processed on different timescales in each of the hemispheres, specifically that the left auditory cortex analyzes information utilizing short (~ 20-40 ms) temporal integration windows, while the right samples information over a longer time frame (~ 150-250ms) (Poeppel, 2003;Boemio et al., 2005). Conveniently, these time integration windows are comparable to the varying timescales of speech e.g. quick spectral changes such as formant transitions happen on time scales of 20-40 ms, whereas syllables occur on timescales of 100-200 ms (Rosen, 1992).
Poeppel proposed that specific modulation of ongoing neuronal oscillatory activity in distinct frequency bands in each hemisphere (i.e. gamma in the left and delta-theta in the right auditory cortices (Poeppel, 2003; Luo and Poeppel, 2012) would reflect sampling and processing of auditory information on two distinct timescales, that match the two most obvious speech rhythms (Schroeder et al., 2008; Giraud and Poeppel, 2012). This suggestion has been supported by recent neuroimaging studies using non-speech stimuli (Giraud et al., 2007; Luo and Poeppel, 2012). Specifically, Luo and Poeppel (2012) using magnetoencephalography (MEG) showed consistent phase locking of theta band activity over right auditory cortex while phase locking in the gamma band was bilateral to non-speech sounds.

Is it possible that lateralized specialization of the auditory cortices exists in non-human primates too? Knowledge of the existence of such a hemispheric specialization could impact the interpretation of results of past and future auditory experiments involving non-human primates, and would have a great impact on determining the validity of the macaque monkey as a model of human auditory processing. Thus far the results of macaque monkey experiments are ambiguous; while early lesion studies (Heffner and Heffner, 1984; Heffner and Heffner, 1986), behavioral studies using the right ear advantage (Petersen et al., 1978; Hauser and Andersson, 1994; Ghazanfar and Hauser, 2001) and studies using neurophysiological approaches (Poremba et al., 2004; Joly et al., 2012) have indicated a left hemisphere dominance in processing conspecific vocalizations, a consensus of right-hemispheric specialization for spectral selectivity in non-human primates remains elusive (Dewson, III et al., 1970; Poremba et al., 2004; Gil-da-Costa and Hauser, 2006). Interestingly, the study by Joly and colleagues (2012) showed that the left higher order areas of auditory cortex (belt and parabelt) of rhesus macaques preferentially responded to human speech and to a lesser degree monkey calls compared to scrambled sounds, indicating that the left hemisphere dominance reflects the
processing of complex spectrotemporal patterns rather than conspecific vocalizations, as human speech has a more intricate spectrotemporal organization.

Therefore the study described in Chapter 3 of the thesis was aimed at investigating whether there is any evidence for a preferential processing of spectrotemporally complex auditory stimuli in left hemisphere A1. We also tested the theory that left A1 aligns its ongoing neuronal activity preferentially to higher, while right A1 to lower temporal scales as proposed by the Poeppel group. To achieve this, we recorded the neuronal activity of left and right hemisphere A1 simultaneously, while the macaques performed a temporal deviant detection task (for detailed methods see Chapter 3) in which we utilized auditory click trains organized on multiple time-scales (corresponding to delta and gamma frequency) similar to the temporal structure of speech. In the majority of human studies that showed auditory functional lateralization mentioned at the start of this section, the subjects were instructed to actively listen to the stimuli, so it is conceivable that this lateralized functionality is only apparent with engagement in the task. To be able to test this, we also presented the same stimuli in a passive condition, where monkeys were alert, but were not required to respond to deviant stimuli. Our specific hypothesis was that neuronal activity will be better aligned to the temporal structure of stimuli when these are attended, and that we will find a hemispheric lateralization of alignment (entrainment): gamma frequency oscillations will be entrained more precisely by click trains in left, while slower, delta frequency oscillations show superior alignment to stimulus structure in right hemisphere A1 neuronal ensembles. Furthermore, contrary to what we predicted in the case of the frequency discrimination task in Chapter 2, we expect that the phase of entrainment will not be frequency dependent, since in the temporal paradigm of Chapter 3 a broadband stimulus (i.e. a click) will be used rather than pure tones. Thus, we anticipate that the majority of A1 will entrain with its high excitability phase during the temporal task in order to boost
detection of the deviant stimulus (i.e. temporal filtering will occur). In other words, since frequency is not an issue in the temporal task and thus spectral filtering is not required of A1, we theorize that the brain should align oscillations across A1 with the same phase to the stimulus structure in order to enhance its neural representation, given that fine temporal structure is the essential feature in this task.

1.5. OBJECTIVES

The overarching purpose of this dissertation was to examine the contribution of ongoing or spontaneous neuronal oscillations to primary auditory cortex responses during attentive and non-attentive conditions, and the role these oscillations play in modulating auditory information. Three intracortical electrophysiological studies, utilizing laminar profiles of synaptic activity (indexed by CSD analysis) and concomitant firing patterns in local neurons (MUA), were proposed to characterize the dynamics of oscillatory neural processes in A1. In summary, the aims and predictions of each study were:

**Study 1 (Chapter 1)** was performed in order to investigate the contribution of phase reset of ongoing oscillations to auditory responses in adjacent A1 ensembles tuned to different frequencies. We predicted that inhibitory responses to non-preferred frequency tones (sideband inhibition) are largely the result of phase resetting of ongoing oscillations to low excitability phases, while BF tone related responses are mixed evoked-phase reset types and in this case oscillations are reset to high excitability phases.

**Study 2 (Chapter 2)** was conducted to examine the effect of selective attention on the frequency tuning of A1 neuronal ensembles, and to determine whether the modulation of ongoing oscillatory activity via phase reset and entrainment could serve as the mechanism of any attention related changes in tuning. We envisioned that during selective auditory attention,
frequency tuning would be sharpened all along the tonotopic axis of A1. The predicted mechanism of the sharpening is that during trial blocks where the attended tone frequency matches the best frequency (BF) of a given recording site, oscillations would be reset to and entrained to their high excitability phases resulting in larger response amplitudes compared to when the same stimuli are ignored. On the contrary, ongoing neuronal activity would be entrained to its low excitability phase if the frequency of attended tones did not match the BF, resulting in a suppression of responses to attended tones.

**Study 3 (Chapter 3)** was carried out to establish whether similar to humans, there is an asymmetry in the temporal processing of attended spectrotemporally complex stimuli across left and right hemispheres. Based on previous human and non-human primate studies examining lateralization of auditory cortex we predicted an overall greater involvement of the left A1 during the performance of the temporal task used in this study, but a possible dichotomy in the alignment to high vs. low frequency temporal task structure across hemispheres. We also speculated that since we used a stimulus with a broad frequency spectrum in the temporal task, oscillations would entrain with their high excitability phase all across A1.
CHAPTER 1

Dual mechanism of neuronal ensemble inhibition in primary auditory cortex.

O'Connell M.N.¹², Falchier A.¹, McGinnis T.¹, Schroeder C.E.¹³, Lakatos P.¹⁴

¹Cognitive Neuroscience and Schizophrenia Program, Nathan Kline Institute, Orangeburg, New York 10962, ²Program in Cognitive Neuroscience, City College CUNY, New York, New York 10031, ³Department of Psychiatry, Columbia College of Physicians and Surgeons, New York, New York 10032, ⁴Department of Psychiatry, New York University School of Medicine, New York, New York 10016

2.1. SUMMARY

Inhibition plays an essential role in shaping and refining the brain’s representation of sensory stimulus attributes. In primary auditory cortex (A1), so-called “sideband” inhibition helps to sharpen the tuning of local neuronal responses. Several distinct types of anatomical circuitry could underlie sideband inhibition, including direct thalamocortical (TC) afferents, as well as indirect intracortical mechanisms. The goal of the present study was to characterize sideband inhibition in A1 and to determine its mechanism by analyzing laminar profiles of neuronal ensemble activity. Our results indicate that both lemniscal and non-lemniscal TC afferents play a role in inhibitory responses via feed-forward inhibition and oscillatory phase reset respectively. We propose that the dynamic modulation of excitability in A1 due to the phase reset of ongoing oscillations may alter the tuning of local neuronal ensembles and can be regarded as a flexible overlay upon the more obligatory system of lemniscal feed-forward type responses.
2.2. INTRODUCTION

Frequency based encoding is a fundamental feature of the auditory system, starting with the spatial ordering of frequency selectivity along the cochlea and continuing with spatially ordered projections and topographically organized frequency maps even beyond primary cortical areas (Merzenich and Brugge, 1973; Kosaki et al., 1997; Kaas and Hackett, 2000). The prevalence of frequency based maps in the auditory system suggests that the extraction of information contained in the frequency content of auditory stimuli is essential for perception and sensory guided behavior. The fact that non-primary auditory areas surrounding primary cortical fields have degraded topographical frequency representations (Kaas and Hackett, 2000) seems to suggest that primary cortical areas play a crucial role in the frequency based computation of auditory representations, since topographically organized feature maps are thought to enhance the efficiency of feature based computations in sensory systems (Kaas, 1997).

The orderly and progressive spatial arrangement of tone frequency neural representations (tonotopic map) is achieved by activation through direct, spatially organized thalamocortical (TC) inputs from the ventral subdivision of the medial geniculate nucleus (MGNv) of the thalamus (Huang and Winer, 2000; Lee et al., 2004; Liu et al., 2007). However it has been shown that anatomical projections from MGN to auditory cortex are not organized in a simple point-to-point fashion, as A1 neurons receive converging inputs from MGN neurons tuned to several ‘neighboring frequencies’ surrounding their best frequency (BF) (Miller et al., 2001; Lee et al., 2004; Winer and Lee, 2007). This predicts that the frequency tuning of A1 should be less sharp than that of MGN, which is not the case (Creutzfeldt et al., 1980; Miller et al., 2002). Hence it has been proposed that intracortical inhibition functions to sharpen the broader pure-tone evoked excitation in A1 in order to enhance response contrast of the neural
representation of frequency (Shamma and Symmes, 1985; Suga, 1995; Sutter et al., 1999). This proposition has been supported by pharmacological experiments showing that the blocking of cortical GABA-mediated inhibition results in a reduction of frequency selectivity in the auditory cortex (Wang et al., 2000; Foeller et al., 2001). In addition to electrophysiological studies, fMRI (Tanji et al., 2010) and optical imaging studies (Horikawa et al., 1996) also indicate suppression of neural activity in response to tones whose neural frequency representation lies in proximity to the BF in a given region of A1 (sideband inhibition). Recent studies have shown that inhibition in A1 can be modulated depending on task demands (Fritz et al., 2003; Fritz et al., 2005a), and can be plastically modified so that A1 is preferentially responsive to behaviorally relevant stimuli (Recanzone et al., 1993; Galindo-Leon et al., 2009).

Although it is apparent from these studies that sideband inhibition in the auditory cortex is necessary for the adaptive processing of behaviorally-relevant, frequency-specific properties of an acoustic stimulus, its precise mechanism has not yet been established. In A1, as in primary somatosensory (Swadlow, 2002; Cruikshank et al., 2007) and visual cortices (Ferster, 1988; Krukowski and Miller, 2001), two types of inhibition provide potential substrates for this effect: 1) feed-forward inhibition mediated by TC afferents from MGNv centered on the granular layer (Wehr and Zador, 2003; Zhang et al., 2003; Tan et al., 2004; Wehr and Zador, 2005; Wu et al., 2008), and 2) lateral type inhibition mediated by intracortical connections that are weighted towards the extragranular layers (Kurt et al., 2008; Moeller et al., 2010). Besides these lemniscal TC and intracortical routes, non-lemniscal TC afferents (Jones, 1998) have also been implicated in modulating the excitability of neuronal ensembles in A1 (Lakatos et al., 2007; Lakatos et al., 2009). These non-specific inputs target mainly the supragranular layers, and seem to be under top-down control. The suggested mechanism underlying this type of modulation is the reset of ongoing neuronal oscillations (Sayers et al., 1974; Makeig et al.,
2004; Lakatos et al., 2009). Since these oscillations reflect the net fluctuation of excitability in a local neuronal ensemble (Young and Eggermont, 2009), reset to their high excitability phase produces facilitation of responses to coincident auditory input, while reset to their low excitability phase can produce a suppression of auditory responses (Lakatos et al., 2007). While the proposed role of the first two types of inhibition in A1 is the sharpening of response timing and frequency tuning, the role of non-specific TC inputs appears to be the dynamical control of cortical excitability based on bottom up and top down influences, and a matching of cortical oscillatory rhythms to those present in task-relevant stimulus streams (Lakatos et al., 2008; Schroeder and Lakatos, 2009). It is not known whether the frequency content of auditory stimuli can change the sign of oscillatory phase reset, and thus whether modulation of neuronal oscillations plays a role in the frequency tuning of A1.

The purpose of this study was twofold: 1) to characterize inhibitory responses to pure tones in A1 in laminar profiles of neuronal ensemble activity, and 2) to examine whether using laminar recordings we could determine which of the above described three mechanisms could play a role in sideband inhibition. We analyzed laminar current source density (CSD) and multiunit activity (MUA) profiles sampled during multielectrode penetrations of A1 in awake macaque monkeys (Macaca mulatta) in response to different frequency pure tones. CSD profiles coupled with the firing of the local neuronal ensemble (MUA) are an invaluable tool in distinguishing between excitatory and inhibitory conductances underlying field potentials, since they provide a reliable index of the location and direction of transmembrane current flow (Mitzdorf, 1985; Schroeder et al., 1998). Also, CSD analysis provides a sensitive measure of synaptic activity even in cases of subthreshold, modulatory responses like oscillatory phase reset (see above). Because our recordings sample all layers simultaneously, we can define and quantify laminar activation profiles, thus generating evidence regarding the relative
contributions of lemniscal and extralemniscal thalamic inputs, as well as corticocortical inputs
(Schroeder et al., 2003). We found evidence that while specific feedforward TC afferents play a
role in both excitatory and inhibitory neuronal ensemble responses, ongoing oscillatory activity
is also modulated in a frequency specific manner by oscillatory phase reset, and the two types
of responses are independent of each other. This layered modulation of excitability in A1
provides a great opportunity for the top down orchestration of excitability across neuronal
ensembles processing different frequencies, so that this cortical area can serve as a complex
spectro-temporal filter in the processing of behaviorally relevant acoustic information.

2.3. RESULTS

In the present study we analyzed laminar CSD and MUA profiles of responses to pure tones
ranging from 353.5 Hz to 32 kHz in half octave intervals obtained with linear array
multicontact electrodes from 64 sites in 9 awake macaque monkeys. The sites were distributed
evenly along the tonotopic axis of primary auditory cortex, with best frequencies (BF) ranging
from ~0.3 kHz to 32 kHz. In addition to the typical excitatory response to BF tones signaled by
a phasic-tonic increase of cell firing (Figure 1A), similar to previous studies (Sutter et al.,
1999; Lakatos et al., 2005a; Steinschneider et al., 2008), most A1 sites responded with a
suppression of MUA, signaling inhibition, to at least one of the pure tones presented at 60 dB
SPL in our suprathreshold tonotopy paradigm (see experimental procedures). We found that
78% of sites (50/64) showed significant post stimulus (15 – 40 ms) MUA suppression
compared to baseline (-50 – 0 ms) to at least one of the pure tones presented (dependent t-test,
p < 0.01). The sites that did not show significant post-stimulus suppression were excluded from
further analysis.
2.3.1. Properties of MUA Suppression in Inhibitory Sidebands

Since the suppression of post-stimulus MUA was largest in the granular layer in the case of all

Figure 1  Excitatory and inhibitory MUA responses. A) On the left is a schematic of a linear array multielectrode positioned in primary auditory cortex. To the right is the laminar profile of averaged MUA in response to the BF tone in area A1. For quantitative analyses of the MUA response, the electrode contact with the largest MUA was selected (red trace). B) Frequency tuning curves for 3 representative A1 sites: one displaying an upper inhibitory sideband (blue, InhibHi), one a lower inhibitory sideband (red, InhibLo), and a third one both an upper and lower inhibitory sideband (green, InhibHi&Lo). Stars denote responses where MUA response amplitudes differ significantly from baseline. Box plots below show pooled frequencies of BF tones and non-BF tones that resulted in the largest inhibition across all sites that had similar tuning curves (InhibHi (n = 19), InhibLo (n = 17), and InhibHi&Lo (n = 14). C) Averaged MUA responses to BF tones and non-BF tones that resulted in the largest inhibition. The grouping of sites is the same as in B). Colored lines below denote time intervals of significant post-stimulus MUA amplitude changes.
inhibitory type responses to pure tones (Tukey’s test, p < 0.01), we first decided to analyze granular layer MUA responses. As Figure 1B illustrates, we noted that inhibition could occur either in response to tones whose frequency was higher (InhibHi, n = 19), lower (InhibLo, n = 17) or both higher and lower (InhibHi&Lo, n = 14) than the BF of a site. These sites were evenly distributed between monkeys. The sites with inhibitory response to higher frequency tones had a low BF (0.3-4 kHz), while the BF of sites with inhibitory responses to low frequency tones was generally high (8-32 kHz). Sites that responded with inhibition to both low and high frequency tones had an intermediate BF (4-16 kHz). While the frequency difference between BF and tones which resulted in the largest MUA suppression was two octaves on average in the InhibHi (mean = 2 octaves, SD = 0.9), and InhibLo groups (mean = 2.3 octaves, SD = 0.88), the average frequency difference between BF tone and maximal inhibition was about half that in the InhibHi&Lo group, on average 1.1 octaves (SD = 0.59), indicating that sites comprising this group were the most sharply tuned (Fig. 1B). While intriguing, this result could be at least partially due to the relatively sparse sampling of frequencies across seven octaves in our experiments.

Figure 2A shows that there is a significant difference between the pooled best frequencies of the three groups while inhibitory frequencies have considerable overlap. As a consequence, there is also a clear difference of MUA response onset latency to BF tones between the inhibitory groups (Fig 2B), since it has been shown that response onset latency to BF tones depends on their frequency (Mendelson et al., 1997;Kaur et al., 2004;Lakatos et al., 2005a). The InhibLo group (highest BF) has the earliest response onset latency (BF3, mean = 7.7ms, SD = 1.0), the next earliest is the InhibHi&Lo group (BF2, mean = 8.3ms, SD = 0.95) and the group with the longest BF tone related MUA response onset latency is the InhibHi group (BF1, mean = 11.9ms, SD = 2.8). The response onset in this last group occurred
significantly later than BF related MUA onset latencies of the other two groups (Tukey’s test, p < 0.01).

The lower part of Figure 2B shows the pooled granular MUA onset latencies of the inhibitory responses in the four inhibitory sidebands (the InhibHi&Lo group has two inhibitory sidebands). While there was an apparent frequency dependence of response onset latencies similar to the pooled BF responses, onset latencies of the inhibitory sidebands did not differ significantly from one another (Kruskal-Wallis test, p > 0.01), which is not surprising, since the pooled frequencies of the inhibitory sidebands have considerable overlap (Fig. 2A). It is also apparent from Figure 2B that while the onset of excitatory (in response to BF tone) and corresponding inhibitory responses in the InhibLo and InhibHi&Lo groups are significantly different (inhibition occurs significantly later) there is no significant difference between the onset of excitation and inhibition in the InhibHi group. The consequence of this can be seen in Figure 1C on the bottom: while in the former two groups (InhibLo and InhibHi&Lo) there is an excitatory MUA spike (excitatory since in laminar CSD

Figure 2 The frequency dependence of granular MUA response onset. A) Box and whisker plot shows the pooled frequency of tones that resulted in the largest excitatory (BF) and inhibitory responses. Grouping and nomenclature are the same as in Fig. 1B) & 1C). B) Pooled granular MUA onset latencies for the 3 BF groups (top) and the 4 non-BF inhibitory response groups (bottom). Brackets indicate where the onset latency of excitatory and inhibitory MUA response is significantly different across InhibHi, InhibLo, and InhibHi&Lo sites (Tukey’s test, p<0.01).
profiles it was paired with a sink) before the inhibition occurs, in the InhibHi group excitation is completely abolished by concurrently occurring inhibition.

The initial excitatory response to tones that result in significant inhibition is probably due to a spread of activation across BF representations in the cochlea and corresponding TC afferents by the relatively high intensity sounds (60 dB SPL) used here (Aitkin and Webster, 1972); the effect emerges in the higher frequency representations of A1 because of the latency advantage of high frequency excitatory responses compared to low frequency inhibitory ones. This short (about 5 ms) excitatory MUA component was present in 14 out of 17 inhibitory responses in the InhibLo group, 9 out of 14 recordings in the lower inhibitory sideband, and in 5 out of 14 recordings in the upper inhibitory sideband of the InhibHi&Lo group. The initial excitation was absent (0/19) in the InhibHi group. We also determined the timing of maximal granular MUA inhibition among the 4 different inhibitory sub-groupings, which was not significantly different between groups (Kruskal-Wallis test, p > 0.01): it occurred on average at 23.4 ms after stimulus onset.

Since one of the mechanisms that has been proposed to underlie off responses in auditory cortex is a rebound from inhibition, we decided to analyze off responses following on responses with largest excitation (BF) and inhibition (non-BF). The rationale was that if this was the case, off responses should be absent following excitatory responses to BF tones, and we should find an off response following most inhibitory responses. First we determined whether granular layer MUA was significantly above baseline following the offset of pure tones in the 115-140 ms time interval (15-40 ms post-offset), and next we verified that the above baseline MUA is not due to a “lingering” of the sustained on response, but rather is a consequence of a phasic MUA increase. We found that 18 out of 50 BF responses (36%) and 22 out of 64 non-BF responses (34%) were followed by an excitatory off response. This
indicates that rather than being merely a consequence of the on response, off responses might be driven by inputs that are distinct from the ones activated by sound onset, with mostly nonoverlapping tuning, as suggested by a recent study (Scholl et al., 2010).

2.3.2. Laminar Profiles of BF vs. non-BF Inhibitory Responses in A1

After functionally identifying cortical layers in each experiment, we compared MUA and CSD laminar response profiles to BF tones and to tones that resulted in the largest inhibition (Fig. 3). While BF tones in all 3 groups resulted in a significant MUA increase in all layers, inhibitory responses were signaled by decreased MUA (examples of a low and high BF site (from the InhibHi and InhibLo groups respectively) are shown in Fig. 3A). The post-stimulus MUA

Figure 3  MUA and current source density (CSD) profiles associated with BF and non-BF tones in A1. A) MUA response profiles from a representative low BF (0.5 kHz, top) and high BF (16 kHz, bottom) site. MUA response profiles to non-BF tones (1.4 kHz and 4 kHz respectively) in the same sites display suppression which is largest in the granular layers. Note that inhibition seems to be fluctuating (two ‘inhibitory’ peaks, one at response onset and one around 100 ms). B) Concomitant CSD response profiles. Note that the laminar pattern of sink source pairs is reversed in response to non-BF tones when compared to BF responses. White arrows shows the supragranular channels with the largest amplitude sink in response to BF tones that were selected for further analyses in these sites.
change compared to baseline was largest in the granular layers in both cases. Analysis of the corresponding CSD profiles in all experiments revealed that as previously described (Schroeder et al., 2003; Lakatos et al., 2007; Steinschneider et al., 2008) the BF tone produces activation of all cortical layers with initial postsynaptic response, a current sink with concomitant increase in neuronal firing in layer 4, followed by later responses in extragranular layers. This sequence of activation, coupled with a source (net outward transmembrane current flow) over sink (net inward transmembrane current flow) in the supragranular layers is typical of an excitatory “feedforward” or “driving” type activation profile (Fig. 3B, left panels).

In contrast to these types of responses, we found that inhibitory responses were of much lower amplitude (note the different scales in Fig. 3B), and that the pattern of sinks and sources in the supragranular layer was inverted in the inhibitory responses compared to excitatory BF responses. The inverted current flow in all cortical layers reflected by the inverted CSD profiles compared to excitatory profiles suggests that inhibition to non-BF pure tones occurs in all cortical layers simultaneously. This was characteristic of all 50 A1 sites analyzed; however in some cases (in the InhibLo and InhibHi&Lo groups) the granular source associated with the inhibitory responses was preceded by a short sink concomitant with the excitatory MUA spike preceding MUA suppression (see above). In addition to an inverted sink-source pattern, another key difference between the BF and non-BF response profiles is that, while the onset of excitatory BF responses is significantly earlier in granular than in supragranular layers (Wilcoxon signed rank, p < 0.01) typical of a feedforward type sequential activation of cortical layers, non-BF responses occur at roughly the same time in granular and supragranular layers (Fig. 4). This suggests that the activation of supragranular layers in the case of responses to non-BF tones may not result solely from a hierarchical interlaminar spread of activation. Rather, this type of laminar response onset profile indicates the influence of TC inputs that
2.3.3. The Physiological Mechanism of the Inhibitory CSD Response

The purpose of our next set of analysis was to determine the mechanism of the inhibitory CSD response. We wanted to see if we could establish whether inhibition is dominated by an evoked type response characterized by a stimulus related CSD amplitude increase in single trials as in the case of BF responses (Lakatos et al., 2007; Lakatos et al., 2009), or a reorganization of ongoing oscillatory activity (phase reset). As a first step, we calculated single trial laminar CSD analytic amplitudes and averaged them across trials. We found that while as expected, there was a significant CSD amplitude increase in the response to BF tones, there was no post-stimulus CSD amplitude increase related to inhibitory responses in any of the layers (Fig. 5, same examples as in Fig. 3B), despite the organized sink-source pattern apparent in the averaged CSD profiles (Fig. 3B).
Post-stimulus CSD amplitude increase was significantly smaller in response to non-BF tones across all sites compared to CSD amplitude increase in response to BF tones (inset in Fig. 5).

Statistical comparison of pre- to post-stimulus CSD amplitudes (averaged across all layers) showed a significant increase in the case of BF responses in all sites (dependent $t$-test, $p < 0.01$), while in most cases no significant effects were found in the case of inhibitory responses (0/19 in InhibHi, 3/17 in InhibLo and 1/14 and 4/14 in the InhibHi&Lo group). This suggests that the mechanism of inhibitory responses seen in the averaged laminar profiles (Fig. 3B) is a modulation of ongoing neuronal activity, rather than increased net transmembrane current flow like in response to BF tones.

If indeed the mechanism of inhibitory responses is the reset of ongoing oscillatory activity as we suspect, this should be reflected by an increased post-stimulus phase coherence across trials (indexed by intertrial coherence – ITC) in frequency bands corresponding to the dominant ongoing oscillations in A1 (Lakatos et al., 2005b) in the absence of pre- to post-stimulus increase in CSD amplitude (Sayers et al., 1974; Makeig et al., 2004). From this point on we do not differentiate between the three inhibitory groups in the description of results, as none of the analyzed variables differed significantly between the groups ($p > 0.01$, Kruskal-
Wallis test). Also, since CSD response amplitude was largest in the supragranular layers, and ongoing/stimulus related oscillations appear to be coherent across cortical layers under a wide range of experimental conditions (Lakatos et al., 2005b; Lakatos et al., 2008), we selected the supragranular electrode channel with largest post-stimulus activity for further analysis (white arrows in Fig. 3B). This selection is also justified by earlier findings showing that the excitability of a cortical column can be reliably linked to CSD oscillations in the supragranular layers (Lakatos et al., 2005b; Lakatos et al., 2007; Lakatos et al., 2008).

To determine whether there was a stimulus related amplitude change in any of the frequency bands, following wavelet decomposition, amplitudes of the single trial responses were computed in several frequency bands ranging from 1 to 100 Hz. Time frequency maps in Figure 6A show an example for excitatory (BF) and inhibitory responses from a representative site. It is apparent that the BF tone causes a large amplitude increase across the entire spectrum except for the low delta frequencies, characteristic of an evoked type complex waveform. On the contrary, we did not find a non-BF tone related amplitude increase in any of the frequency bands; the poststimulus amplitude trace is almost an exact match to the prestimulus one. Comparison of the pre- and post-stimulus oscillatory amplitudes in 6 frequency bands (1-2.4 Hz; 2.4-4 Hz; 4-10 Hz; 10-15 Hz; 15-25 Hz; 25-60 Hz) revealed no significant post-stimulus amplitude change for any of the inhibitory responses (dependent $t$-test, $p < 0.01$).

Next we calculated intertrial coherence (ITC) values across the single trials in each experiment, to determine whether the inhibitory laminar pattern seen in the average CSD profiles was due to an event-related phase synchrony across trials. The value of ITC will be 1 in the extreme case if the oscillatory phase is the same in each trial, and it will be 0 if the oscillatory phase across trials is random. Figure 6B shows that while the supragranular response to the BF tone is characterized by high ITC values across a wide range of frequencies,
typical of an evoked type waveform (Lakatos et al., 2007; Lakatos et al., 2009), the inhibitory response is associated with phase locking (non-random phase distribution across trials indexed by higher ITC values) in three distinct frequency bands, which are the dominant oscillations present in the ongoing (prestimulus) neuronal activity (Fig. 6A and Lakatos et al., 2005b; Lakatos et al., 2007). This provides further evidence that phase reset perturbs the phase of ongoing activity without radically changing its overall composition.

The poststimulus ITC peaks in the delta (1-4Hz), theta (4-10 Hz) and gamma (25-55 Hz) frequency ranges were detectable in all inhibitory responses, and statistical testing showed
that they signaled non-random phase distribution in all cases (Rayleigh p < 0.01). The mean ITC value was 0.49 (SD = 0.18) in the delta, 0.40 (SD = 0.14) in the theta and 0.34 (SD = 0.14) in the gamma band for 100 single trials on average (variation between number of trials across experiments was relatively small, SD = 9.2). As the boxplots in Figure 6B illustrate, there was no difference between the frequencies of ITC peaks in the delta, theta and gamma ranges between responses to BF tones and inhibitory responses (p > 0.01, Wilcoxon signed-rank test), which indicates that both types of responses result in the phase reset of dominant ongoing oscillations, even though this cannot be unambiguously shown in the case of BF responses because of the evoked type response that results in wideband ITC even in the absence of phase reset (Lakatos et al., 2009). Taken together, the above results show that responses to BF tones are mixed - evoked and phase reset - type, while the mechanism of inhibitory responses is predominantly oscillatory phase reset.

It is well documented that neuronal oscillations reflect rhythmic changes of excitability in neuronal ensembles (Young and Eggermont, 2009). Thus, if the phase ongoing oscillations are reset to is dependent on the frequency of auditory stimuli, this would aid in sharpening the tuning of different frequency regions in A1. To examine this possibility, we decided to compare the post stimulus phase of reset oscillations in the dominant frequency bands for both BF tone related and inhibitory responses. To determine the post-stimulus time instant at which to evaluate oscillatory phases, we calculated the mean timing of the maximum gamma ITC peak across inhibitory responses (mean = 23.4 ms, SD = 6.6). Interestingly, we found that it was not significantly different from the timing of the mean maximal MUA inhibition (23.46 ms, see above).

Histograms in Figure 6C show post stimulus single trial phase distributions for delta, theta and gamma oscillations associated with excitatory (left) and inhibitory (right) responses
for a representative recording site. In both response types, there is a clearly significant grouping of phases in each frequency band. It is also apparent that the mean phases of the post-stimulus oscillations (black dotted line on histogram) are significantly different. In fact oscillations in the BF and non-BF conditions are in counter-phase, with the exception of the delta band. In the case of BF tone stimulation, the phases are grouped before and around the negative peak of the oscillations (±π in Fig. 6C&D), which has been shown to correspond to the high excitability phase of ongoing CSD oscillations at this laminar location (Lakatos et al., 2005b; Lakatos et al., 2007). Conversely, single trial oscillatory phases in the inhibitory response are clustered around the positive peak which corresponds to the low excitability phase of ongoing oscillations at this supragranular location. Figure 6D displays the distribution of pooled mean phases associated with excitatory (n = 50) and inhibitory responses (n = 64), which shows a similar pattern. There is significant non-uniform phase distribution of the mean phases in each frequency band in the case of responses related to BF tones around the high excitability phase (negative peak) of ongoing oscillations (Rayleigh’s uniformity tests, p < 0.01). In the case of inhibitory responses only the mean theta phases show a significantly non-uniform distribution around the low-excitability phase (opposite to that in the excitatory responses). While there is also apparent grouping in the distribution of gamma phases opposite to gamma phases in the excitatory response, this did not prove to be significant (Rayleigh’s uniformity tests, p = 0.06), possibly as a consequence of response onset variation.

Delta oscillations are somewhat special compared to higher frequency oscillations in these experiments, since their frequency overlaps the frequency of stimulus presentation in our paradigms (SOA = 624.5 ms corresponding to a presentation frequency of 1.6 Hz). This means that they can entrain to the presentation rate of auditory stimuli used in our suprathreshold tonotopy paradigm as we showed previously (Lakatos et al., 2005b). To complicate matters,
while in most experiments we used a tonotopy paradigm that consisted of a random stream of different frequency pure tones, in 14 of our 50 (28%) penetrations we delivered each frequency tone to the subjects in a blocked design. Could this be a source of the seemingly random phase distribution observed in Figure 6D?

Figure 7A shows the mean delta phase distributions of excitatory and inhibitory responses for blocked and random streams of pure tones. It is apparent that in the blocked design, mean delta phases associated with inhibitory non-BF tones are opposite to mean delta phases associated with streams of BF tones, and are clustered around the low excitability phase. In contrast, in the remaining 36 experiments where pure tones were delivered in a random order, delta oscillations associated with BF and inhibitory stimuli both entrained to the high excitability phases in most cases (Fig 7A, right). The mean of the mean delta phases (black dotted lines in Figure 7A) were significantly different between excitatory and inhibitory responses in the blocked condition (Fisher’s nonparametric test for the equality of circular means, p < 0.01).
This means that, as the averaged supragranular waveforms in Figure 7B illustrate, in the case when stimulus streams are formed by pure tones of the same frequency, supragranular delta oscillations can entrain to BF and non-BF stimulus streams with opposing phases. However, if the rhythmic stimulus stream consists of unpredictable frequency stimuli in a wide frequency range, delta oscillations tend to entrain to the stimulus stream with their high excitability phases. Thus it seems likely that in the case of rhythmic stimuli that consist of a narrow frequency range, entrained delta oscillations act as rhythmical “filters” by enhancing responses to BF and suppressing responses to non-BF tones. We have to note, that the experiments that yielded the data for the present study were not specifically designed to distinguish between the effects of uniform and random frequency content in rhythmic stimulus streams on oscillatory entrainment. Further studies that parametrically vary the broadness of the frequency content and stimulation rate are needed to study the dynamics of this effect, and to decide whether delta oscillations are special in this regard.

Our finding that a phase difference (opposition) of neuronal oscillations in local neuronal ensembles can occur at distances as small as 1 mm – which corresponds to approximately an octave difference in tuning – across all the frequency ranges of the oscillatory spectrum investigated in the present study indicates that even low frequency (delta) oscillations can be fairly local. To further specify the characteristics of local neuronal oscillatory activity, one would have to record ongoing and event related neuronal activity at varying distances simultaneously across A1. We speculate that while ongoing oscillatory activity may be independent even at the scale of neighboring cortical columns (dimensions well under 0.5 mm), the independence of event related (phase reset) oscillatory activity is restricted by the selectivity of non-specific thalamocortical inputs, which are known to project more widely than lemniscal inputs.
2.4. DISCUSSION

In the present study, we found that about 80% of A1 sites respond with significant MUA suppression to pure tones with frequencies that differ from their preferred frequency (BF). Analysis of MUA and concomitant CSD laminar response profiles in these sites revealed that BF tones produced a strong activation of all cortical layers with an initial postsynaptic response (current sink with concomitant increase in action potentials) in the granular layer, followed by later responses in extragranular layers, typical of an excitatory “feedforward” or “driving” type activation profile (Schroeder et al., 2003; Lakatos et al., 2007). On the other hand non-BF inhibitory tones produced a weaker CSD response characterized by a simultaneous activation of all cortical layers, suggesting – similar to granular layers – a direct activation of supragranular layers. In addition, the arrangement of sinks and sources was inverted in the laminar profiles of inhibitory responses relative to BF responses. Single trial time frequency analysis of the event-related oscillations revealed that – at least in the supragranular layers – responses associated with non-BF inhibitory tones were most likely a result of phase resetting of ongoing oscillations within specific frequency bands. We also found that while in the mixed evoked-phase reset type responses to BF tones, oscillations are reset to high excitability phases, in response to non-BF tones, ongoing oscillatory activity is generally reset to opposing, low excitability phases.

2.4.1. Prevalence of Inhibition in Primary Auditory Cortex

We found that 14 out of 64 A1 sites (22%) investigated in the present study did not show a significant post stimulus MUA suppression to any of the pure tones presented. Even though we did not find a significant post-stimulus inhibition in these sites, responses to pure tones whose
frequency was about 2 octaves away from the BF showed the smallest, in most cases slightly below baseline post-stimulus MUA. A potential reason why we did not find MUA suppression in all A1 sites investigated is that attention was not engaged during the presentation of the stimuli. Since it has been shown that the phase reset of ongoing oscillatory activity is strongly dependent on stimulus salience (Lakatos et al., 2009), it is likely that if auditory stimuli were made task-relevant, this would result in stronger phase reset. Our finding that the phase to which ongoing oscillations are reset to depends on the stimulus frequency predicts that both inhibition and excitation would be facilitated under attentive listening conditions, resulting in sharper tuning. Since in our experiments attention was not directly manipulated, it remains to be tested in behaving animals attending to specific frequencies whether the common finding that frequency tuning in auditory cortex can be modified by attention (Fritz et al., 2003; Fritz et al., 2005a; Bidet-Caulet et al., 2007; Kauramaki et al., 2007; Okamoto et al., 2007) is due to this mechanism.

Similar to prior studies, our results demonstrate that inhibitory sideband asymmetry depends on the BF, which was suggested to underlie the topographic organization of FM direction selectivity in primary auditory cortex (Zhang et al., 2003). In addition to an opposite direction sideband asymmetry, another difference between sites with low and high BFs is that while in sites with low BFs, inhibition to high frequency sounds is “complete”, sites with high BF tend to respond with a short excitation to low frequency sounds before the inhibitory response. This could be simply an ‘artifact’ of the different response onset latencies related to low and high frequency stimuli (discussed below), but it could also serve an important role in the perception of spectrally complex stimuli, like species specific communication or speech. If true, this would indicate that similar to functional differences along the isofrequency axis
Middlebrooks et al., 1980; Cheung et al., 2001; Read et al., 2002), functional differences also exist along the tonotopic axis in primary auditory cortex.

2.4.2. Mechanisms of Inhibitory Responses

Our results show that while the excitation driven by the BF tone occurs earlier in the granular than in the supragranular layers indicating a sequential laminar activation, non BF tone-related inhibition occurs simultaneously across the layers, suggesting parallel activation. This indicates that – at least initially – granular and supragranular inhibition might occur via different routes and mechanisms, which we will first discuss separately beginning with the granular layer.

We found that similar to excitatory responses, inhibitory response onsets to pure tones also depend on their frequency, and that in general, inhibitory response onset to a given pure tone is about 3-4 ms later than excitatory response to the same frequency tone in the granular layer in both MUA and CSD profiles (Figures 2 & 4), which parallels the delay between excitation and inhibition reported by Wehr and Zador, (2003). This suggests that the mechanism of initial granular layer MUA suppression is feedforward inhibition via specific TC afferents, and the delay is due to the disynaptic nature of the inhibition, as opposed to the monosynaptic excitatory response. (Swadlow, 2002; Wehr and Zador, 2003; Zhang et al., 2003; Tan et al., 2004; Cruikshank et al., 2007; Wu et al., 2008).

How can this mechanism explain that while – as expected based on the delay of inhibition compared to excitation – there is always a short excitation preceding inhibition at high frequency A1 sites, inhibition to high frequency tones in low frequency sites completely abolishes excitation? A key observation in this regard is that the difference between the onset of excitation and suppression is on the scale of the difference between the onset of responses to different frequency BF tones. Another pre-requisite for this finding is that excitation and
inhibition are mediated by TC afferents from different frequency regions of the MGNv. One way this is possible is if inhibitory receptive fields in A1 are broader than excitatory ones, like in primary somatosensory cortex (Swadlow, 2002). High intensity pure tones, like the ones used in the present study will activate a considerably broad area of the cochlear receptor surface around the region corresponding to the pure tone frequency, which will activate TC afferents in a relatively wide frequency band (Aitkin and Webster, 1972). While specific feedforward (layer 4) activation of the excitatory neurons in a given area is mediated by TC neurons that match the BF of the site, suppression is mediated through TC afferents tuned to other frequencies as well. Therefore, if suppression is activated by TC afferents that are tuned to higher frequencies and thus are activated faster, this suppression can completely prevent weaker excitatory responses mediated by lower frequency TC afferents that are ‘slower’.

To summarize, our results confirm that feedforward inhibition in primary auditory cortex is more broadly tuned than excitation (Wu et al., 2008), and this results in characteristic interactions of excitation and inhibition in the neuronal ensemble responses to non-BF tones that is dependent on the frequency relation of the non-BF tone to the BF. In theory, the broader tuning of inhibitory cell populations could be mediated by a slightly different, more divergent set of TC inputs than those mediating excitation, similar to the suggested TC connectivity of the barrel cortex (Swadlow, 2002), however a recent study in primary auditory cortex found that the frequency range of TC inputs is similar between excitatory (regular spiking) and inhibitory (fast spiking) neurons (Wu et al., 2008). Thus, the broader tuning of feedforward inhibition that results in the lateral sharpening of frequency tuning is likely due to less selective outputs: inhibitory neurons are capable of converting a broader range of synaptic input into action potentials than excitatory ones (Cruikshank et al., 2007).
In a recent study, Atencio and Schreiner (Atencio and Schreiner, 2010) found that interlaminar differences in temporal and spectral modulation transfer functions in A1 cannot be explained by a purely sequential interlaminar flow of information, thus suggesting the influence of non-lemniscal thalamocortical and/or horizontal inputs on auditory stimulus processing in A1 cortical columns. Our results also make the case for this, since we found that response onset latency to non-BF tones is not significantly different in the granular and supragranular layers, thus they are not activated sequentially like in the case of typical feedforward type responses to BF tones. A second potential route would be activation through horizontal intracortical inhibitory (Tomioka et al., 2005) or excitatory (Kurt et al., 2008) connections. However, if we consider, that in cases where inhibition occurs to higher frequency tones than the BF, response onset in the supragranular layers is the same to BF and non-BF tones (InhibHi and InhibHiLo2 in Figure 4), then this route seems unlikely, since it should result in a more significant delay of the non-BF compared to the BF response onset (even if we consider that response onset to a 2 octave higher tone would be approximately 2 ms earlier, resulting in an earlier activation of horizontal fibers). Thus, the most likely candidate is the third possible route which is activation of the supragranular layers by non-specific (non-lemniscal) TC afferents. This “activation” as our study revealed results in the frequency specific reset of ongoing neuronal activity through afferents that most likely originate in the medial region of the MGN, since it has been shown that these “non-specific” thalamic afferents target mainly supragranular neuronal ensembles (Roger and Arnault, 1989;Hashikawa et al., 1991;Molinari et al., 1995;Jones, 1998;Huang and Winer, 2000).

We can only speculate about the mechanism that enables frequency specificity of the phase (high vs. low excitability) ongoing oscillations are reset to. We think that this mechanism is most likely thalamic, for the reason that the thalamus seems to be a better strategic location
to orchestrate the coherent activity of neuronal populations across A1 than an intracortical mechanism. We also speculate that switching between an excitatory and inhibitory type phase reset in response to different frequency tones might involve the reticular nucleus of the thalamus (TRN) since this structure is the major source of inhibition to the MGN (Crabtree, 1998; Guillery et al., 1998). If this would be the case, the rich connections of the TRN with prefrontal cortical areas (Zikopoulos and Barbas, 2006) could explain how top-down influences are able to modulate the strength of oscillatory phase reset (Lakatos et al., 2009).

New techniques, such as the lentivirus mediated expression of photosensitive ion channels (Cruikshank et al., 2010) selectively in non-specific thalamocortically projecting neurons might provide invaluable information about the anatomical substrates of oscillatory phase reset of sensory oscillations in the near future. Selective pharmacological silencing of intracortical activity in a given A1 region (with the GABA\textsubscript{A} receptor agonist muscimol, similar to (Happel et al., 2010) could also help to disentangle the contribution of horizontal and non-specific thalamocortical inputs in responses to non-BF tones. We suggest that to achieve this, one would have to selectively silence a region of A1 where a given pure tone results in evoked type (lemniscal) activity, and record the neuronal activity of an A1 site that is unaffected by the muscimol effect. The reason for this is that a complete muscimol blockade of A1 would eliminate intracortically evoked activity, but would also abolish ongoing neuronal oscillations, thus oscillatory phase reset. In contrast, the selective silencing of the A1 region receiving lemniscal activation in response to a given frequency pure tone would effectively block the spread of “specific activity”, while still enabling modulation of cortical activity through non-specific thalamocortical inputs in regions unaffected by pharmacological manipulation.
It is important to note that while oscillatory phase reset is most prevalent in the supragranular layers, the phase of ongoing supragranular oscillations reflects excitability changes in all laminae of cortical processing units (Lakatos et al., 2005b; Lakatos et al., 2007; Lakatos et al., 2008). This is illustrated by Figure 8, which shows the phase triggered (grey arrows) averages of spontaneous CSD for two A1 sites. To facilitate comparison to the ‘real’ inhibitory responses, these sites are the same as in Figure 3, and we created a ‘baseline’ from randomly selected spontaneous epochs. The corresponding laminar MUA profiles show a remarkable similarity to the ‘real’ inhibitory responses, clearly establishing the possibility that inhibitory responses can emerge as a result of pure phase resetting of neuronal oscillations. Of note is that the MUA ‘suppression’ related to the low excitability phase of ongoing oscillations is largest in the granular layer compared to MUA related to random ongoing activity (baseline), and the MUA amplitude change appears cyclically fluctuating, just as it occurs in the inhibitory responses that we directly measured at these sites. In both cases there are two MUA suppression peaks, one at stimulus onset and one around 100 ms separated by a time period where MUA seems to return to baseline, suggesting

**Figure 8 Phase triggered averages of spontaneous CSD and concomitant recorded MUA.** A) The two phase triggered average profiles were created from spontaneous activity recorded in the same locations as CSD response profiles in Figures 3 & 5. The phase triggered profiles (starting at the arrow) are the average of epochs of ongoing activity triggered at phases of supragranular delta, theta and gamma oscillations that correspond to the mean phase of these oscillations in the inhibitory responses. The “baseline” (activity preceding the arrows) was created from averaging randomly selected epochs of ongoing activity. B) Laminar profiles of concomitant MUA.
a cyclical modulation around 10 Hz, which roughly corresponds to the wavelength of the dominant theta oscillation that is reset by auditory inputs.

These considerations suggest that while – for reasons detailed above – there is undoubtedly an initial inhibitory component related to specific feedforward TC pathways, the bulk of the inhibitory response, especially that following the early part of the response, is due to the phase reset of ongoing supragranular oscillations in A1. Selective blockade of non-specific thalamocortical inputs could provide definitive proof for this hypothesis. Note that even though we used the lower supragranular electrode site to ‘trigger’ epochs of spontaneous activity at specific phases of ongoing (Fig. 8), a sink-source pattern can be seen throughout all cortical laminae, indicating that CSD activity across the cortex is – at least to some degree – coherent or coupled to each other, as suggested by previous studies (Sakata and Harris, 2009). Even the related MUA across all layers is either enhanced or suppressed simultaneously. Resetting cortical oscillations to a low excitability phase (present data) can aid specific feed forward inhibition and temporally extend its effects by lowering the membrane potential of excitatory neuronal populations, thereby tilting the balance of concurrently occurring specific TC input generated excitatory/inhibitory processes towards inhibition, and effectively preventing an excitatory response.

The intriguing finding that delta oscillations entrained differently to rhythmically presented stimulus streams based on the composition of the streams (narrow vs. wideband frequency content) indicates that concurrent with stimulus selection (Lakatos et al., 2008), entrained slow oscillations might play an important role in auditory stream segregation. Our results indicate the reset and entrainment of ongoing oscillatory activity by auditory inputs can be frequency specific, we propose that neuronal oscillations in A1 can “track” frequency and timing in attended auditory streams simultaneously. By arranging high excitability phases to
coincide with key events in attended stimulus streams in both frequency (across A1) and time, attended streams get amplified and segregated along the two arguably most fundamental organizing dimensions in auditory processing. As a “bonus”, streams that do not closely match either in frequency or time get suppressed by the low excitability phase of the entrained oscillations. To achieve this, oscillations would have to be simultaneously orchestrated (via phase reset and entrainment) across A1 by frequency specific inputs, which could be verified by multiple site recordings across A1.

2.4.3. Conclusions

Our findings outline a dual mechanism of inhibition in A1. While one mechanism is mediated by specific, lemniscal thalamocortical inputs targeting layer 4, the other involves non-specific thalamocortical inputs targeting the supragranular layers. This latter mechanism involving the phase reset of ongoing oscillations is more dynamic and, based on earlier studies, has the potential to change the strength and possibly even shift the tuning of local neuronal ensembles in A1. Along with modulating excitability locally, this dynamic overlay of ongoing oscillatory activity is an ideal candidate for the orchestration of neuronal activity across A1 when processing complex auditory scenes.

2.5. EXPERIMENTAL PROCEDURES

2.5.1. Subjects. We analysed electrophysiological data recorded during 64 penetrations of area A1 of the auditory cortex of 1 female and 8 male macaques (*Macaca mulatta*) weighing 5-9 kg, who had been prepared surgically for chronic awake electrophysiological recordings. No monkeys were used exclusively for this study; rather, they were all assigned to other primary experiments. Because all of our auditory cortex experiments require functional identification of
recording sites using a battery of pure tone and broadband noise stimuli (see below), the data generated by the routine methodological procedures were available for the analyses outlined below. Prior to surgery, each animal was adapted to a custom fitted primate chair and to the recording chamber. All procedures were approved in advance by the Animal Care and Use Committee of the Nathan Kline Institute.

2.5.2. Surgery. Preparation of subjects for chronic awake intracortical recording was performed using aseptic techniques, under general anesthesia, as described previously (Schroeder et al., 1998; Mehta et al., 2000). The tissue overlying the calvarium was resected and appropriate portions of the cranium were removed. The neocortex and overlying dura were left intact. To provide access to the brain and to promote an orderly pattern of sampling across the surface of the auditory areas, matrices of 18 gauge stainless steel guide tubes or plastic recording chambers (Crist Instruments) were positioned normal to the cortical surface of targeted areas for orthogonal penetration of area A1 in the superior temporal plane. These matrices were angled so that the electrode track would be perpendicular to the plane of auditory cortex, as determined by pre-implant MRI. Individual epidural guide tubes were positioned over central and frontal sites to serve as ground and reference electrodes. Together with socketed Plexiglas bars (to permit painless head restraint), they were secured to the skull with orthopedic screws and embedded in dental acrylic. A recovery time of two weeks was allowed before the beginning of data collection.

2.5.3. Electrophysiology. Animals sat in a primate chair in a dark, isolated, electrically shielded, sound-attenuated chamber with head fixed in position, and were monitored with infrared cameras. Laminar profiles of field potentials (EEG) and concomitant population action potentials (multiunit activity or MUA) were obtained using linear array multi-contact electrodes (24 contacts, 100 μm intercontact spacing). The multielectrode was inserted acutely
through a guide tube (either implanted permanently or inserted in the recording chamber for the time of recording) sited above the area of interest for that session; it was lowered through the dura into the brain, and positioned with the electrode channels spanning all layers of the cortex (Fig. 1A). Signals were impedance matched with a pre-amplifier (10x gain, bandpass dc-10 kHz) situated on the electrode, and after further amplification (500x) the signal was split into field potential (0.1-500Hz) and MUA (300-5000Hz) range by analogue filtering. Field potentials were sampled at 2 kHz/16bit precision; MUA was sampled at 20 kHz/12bit precision. Additional zero phase shift digital filtering (300-5000Hz) and rectification was applied on the MUA data to extract the continuous estimate of cell firing. One-dimensional current source density (CSD) profiles were calculated from the local field potential profiles using a three-point formula for estimation of the second spatial derivative of voltage (Nicholson and Freeman, 1975; Schroeder et al., 1998). CSD profiles provide an index of the location, direction, and density of the net transmembrane current flow, the first-order neuronal response to synaptic input (Mitzdorf, 1985; Schroeder et al., 1998). During each experiment, the laminar CSD profile evoked by binaural broadband noise bursts (BBN) was used to position the multielectrode array to straddle the auditory cortex from the pial surface to the white matter (Schroeder et al., 2001). At the beginning of each experimental session, after refining the electrode position in the neocortex, we established the best frequency (BF) of the recording site using a “suprathreshold” method (Steinschneider et al., 1995; Schroeder et al., 2001; Fu et al., 2004). The method entailed either the presentation of 7 different frequency pure tones ranging from 500 to 32000 kHz in one octave steps in separate blocks (14 out of 50 experiments), or a pseudorandom train of 14 different frequency pure tones ranging from 353.5Hz to 32kHz in half octave steps, and a broadband noise burst (BBN) at 60 dB SPL (duration: 100 ms, r/f time: 4 ms, ISI = 767, n = 1000). Auditory stimuli were produced using Tucker Davis Technology’s
System III coupled with ES-1 speakers. Both the resulting field potential and MUA were stored as continuous records.

2.5.4. **Data analysis.** Data were analyzed offline using Matlab (Mathworks, Natick, MA).

After selective averaging of the CSD and MUA responses to the fourteen randomly presented pure tones and BBN, we determined the best frequency (BF) of the recording site. Recording sites were functionally defined as belonging to AI or belt auditory cortices based on the sharpness of frequency tuning, the inspection of the tonotopic progression across adjacent sites, and relative sensitivity to pure tones versus broad-band noise of equivalent intensity (Merzenich and Brugge, 1973; Rauschecker et al., 1997; Schroeder et al., 2001). In the present study only recordings obtained from area A1 were analysed. At the end of each animal’s experimental participation, functional assignment of sites was confirmed histologically (Schroeder et al., 2001; Fu et al., 2004).

Utilizing the BF-tone related laminar CSD profile, the functional identification of the supragranular, granular and infragranular cortical layers in area A1 is straightforward based on our earlier studies (see Fig. 3B) (Schroeder et al., 1998; Schroeder et al., 2001; Fu et al., 2004; Lakatos et al., 2005a; Lakatos et al., 2007). For quantitative analysis of event related MUA amplitudes, the electrode contact with the largest BF tone-related MUA was selected, which was found to always reside in the granular layer (red trace in Fig. 1A). To determine if a site displayed significant inhibition, single trial mean MUA amplitudes were calculated for the time window 15-40ms on this channel (the transient part of the responses (Steinschneider et al., 2008), and compared to the baseline (-50 to 0ms) (dependent t-test, p < 0.01). For our initial analyses we divided the A1 sites into groups based on the relationship of the tones that showed maximal inhibition (non-BF inhibitory tone; tone that evoked the largest granular MUA inhibition) to the best frequency tone (BF) (Fig. 1B). To determine MUA response onset
latencies, the same granular electrode was used and response onset was defined as the earliest significant (> 2 standard deviation units) deviation of the averaged waveforms from their baseline (-50–0 ms), that was maintained for at least 5 ms. In the case of inhibitory responses, the onset of inhibition was determined by only taking negative direction MUA changes into account. Pooled onset latency values were evaluated statistically using Kruskal-Wallis test (p < 0.01) and post hoc Tukey’s test across the different groups (see Fig. 2B). The analysis of CSD onset latencies (Fig. 4) was performed similarly on two selected channels, one from supragranular and one from the granular layers. These electrode channels were selected based on which site had the largest amplitude sink in these layers in response to BF tones.

To extract auditory event related CSD amplitudes (Fig.5), we calculated the analytic amplitude of the single trial CSD signals for the entire pass-band using the Hilbert transform. To statistically evaluate whether stimulus related responses resulted in a difference between pre- and post-stimulus CSD amplitude, we averaged the single trial analytic CSD amplitude across all cortical layers, and then compared these variables averaged in the pre-stimulus (-50-0 ms) and post-stimulus (0-100 ms) time intervals within experiments using dependent t-tests (p < 0.01).

For the analysis of event related CSD oscillations, continuous recordings were epoched from -2000 to 2000 ms to avoid edge effects of the wavelet transformation. After selecting the supragranular electrode with the largest amplitude BF-tone related sink, instantaneous power and phase were extracted by wavelet decomposition (Morlet wavelet) on 84 scales from 1 to 101.2 Hz for single trials in response to BF and non-BF tones. To determine stimulus related oscillatory amplitude changes, we statistically compared pre- (-100 to 0ms) and post-stimulus (0-100ms) oscillatory amplitudes (dependent t-test, p < 0.01) in 6 frequency bands (gamma (25-55), beta (13-25), alpha (10-13), theta (4-10), high delta (2-4 Hz), and low delta (1-2 Hz).
Frequency limits were chosen based on results from our earlier studies (Lakatos et al., 2005b; Lakatos et al., 2007). To characterize the phase distribution across trials, the wavelet transformed single trial data was normalized (unit vectors), trials were averaged, and the length (modulus) of the resulting vector was computed (see e.g. Lakatos et al., 2007). The mean resultant length, also called inter-trial coherence (ITC) ranges from 0 to 1; higher values indicate that the observations (oscillatory phase at a given time-point across trials) are clustered more closely around the mean than lower values. Single trial event-related phase values were analyzed by circular statistics. If non-random phase distribution is related to an event, it is also called phase locking. Significant phase locking - deviation from uniform (random) phase distribution - was tested with Rayleigh’s uniformity test. Visual inspection of ITC spectrograms revealed peaks in three distinct frequency bands; the low delta (1-2 Hz), theta (4-10 Hz) and gamma (25-55 Hz) frequency ranges (Fig. 6B). The peak ITC values and the time of their occurrence (in the 0-100 ms post-stimulus time interval) in these bands were selected automatically in most cases by searching for maxima within a given frequency range and the 0-100 ms time frame. Pooled phase distributions (Fig. 6D) were compared by a nonparametric test for the equality of circular means (Fisher, 1993; Rizzuto et al., 2006). The α value was set at 0.01 for all statistical tests.

To create the phase triggered averages of spontaneous CSD and concomitant recorded MUA in Figure 8, in non-overlapping 2 second segments of spontaneous activity we determined the time points that most closely corresponded to the combination of mean delta, theta and gamma oscillatory phases measured in real inhibitory responses (e.g. Fig 6C). To achieve this, we found the closest phase value to the mean delta phase in the 2 second segment, then the theta phase closest in time to this time point that corresponded to the mean theta phase, and finally the gamma phase closest to this time point that corresponded to mean gamma phase
in inhibitory responses. Using these time-points as ‘triggers’, we created 100 epochs starting at these time-points (arrows in Figure 8) from the spontaneous CSD and concomitant MUA. To simulate a baseline where oscillatory phase is random, we also created 100 epochs ending at randomly selected time-points from the same spontaneous CSD and MUA recordings, and merged these epochs with the phase triggered ones.

Besides phase analyses, the entrainment of delta oscillations was verified visually in averaged surpapranular waveforms, since cyclically occurring evoked responses can artificially bias the phase of oscillations at the frequency of stimulus presentation. As far as we know, the only way to distinguish between real entrainment and the artificial phase bias caused by repetitive evoked type waveforms is a visual inspection of the averaged waveforms: if the mean single trial phase matches the direction of fluctuation we observe in the baseline, we can talk about entrainment.
CHAPTER 2

Sharpening of frequency tuning by selective attention in primary auditory cortex.

O’Connell M.N.1, 2, Barczak A.1, 4, Schroeder C.E.1, 2, 3, Lakatos P. 1, 3

1Cognitive Neuroscience and Schizophrenia Program, Nathan Kline Institute, Orangeburg, New York 10962, 2Program in Cognitive Neuroscience, City College CUNY, New York, New York 10031, 3Department of Psychiatry, Columbia College of Physicians and Surgeons, New York, New York 10032, 4Department of Neuroscience & Physiology, New York University School of Medicine, New York NY
3.1. SUMMARY

Recent electrophysiological and neuroimaging studies provide convincing evidence that attending to sounds increases the response selectivity of neuronal ensembles even at the first cortical stage of auditory stimulus processing in primary auditory cortex (A1). This is achieved by enhancement of responses in the regions that process attended frequency content, and by suppression of responses in the surrounding regions. The main goals of our study were to define the extent to which A1 neuronal ensembles are involved in this process, determine its effect on the frequency tuning of A1 neuronal ensembles, and examine the involvement of the different cortical layers. To resolve these issues, we analyzed laminar profiles of synaptic activity and action potentials in area A1 of non-human primates performing a rhythmic intermodal selective attention task. We found that frequency tuning was sharpened all along the tonotopic axis of A1 due to both increased gain at the preferentially processed or best frequency and increased response suppression at all other frequencies. Our results suggest these effects are due to a frequency-specific entrainment of ongoing delta oscillations, which orchestrates frequency-specific predictive excitability changes across all of A1. This results in a net suppressive effect due to the large ratio of neuronal ensembles that do not specifically process the attended frequency content. In addition, analysis of the laminar activation profiles revealed that while attention related suppressive effects predominate the responses of supragranular layer neurons, response enhancement is dominant in the granular and infragranular layers - providing evidence for layer specific cortical operations in attentive stimulus processing.
3.2. INTRODUCTION

Although it has been well over 100 years since William James (1890) stated that “everyone knows what attention is”, its underlying neural mechanisms are still being debated today. It is still unclear, for example, to what extent auditory attention increases response gain (resulting in increased neural activity) and/or response selectivity (resulting in sharper tuning), and how these attentional neuronal effects are achieved. Early studies of visual attention using diverse methodologies have indicated the existence of a sensory response amplification or gain mechanism for attended compared to ignored locations or features (Harter et al., 1982; Hillyard, 1984; Corbetta et al., 1990; Heinze et al., 1994; O'Leary et al., 1996; Luck et al., 1997; Lakatos et al., 2008). Results of more recent studies have supplemented this view by suggesting that attention to specific visual features may also enhance response selectivity or response contrast of visual neuronal populations (Murray and Wojciulik, 2004; Womelsdorf et al., 2006a; Fischer and Whitney, 2009). However this point is still under debate as findings from single unit neurophysiology experiments in non-human primates conflict on this point (Spitzer et al., 1988; Desimone and Duncan, 1995; McAdams and Maunsell, 1999; Treue and Maunsell, 1999; Treue and Martinez Trujillo, 1999).

Mechanistic, animal-model studies of attention in the auditory system are far less common than in the visual system, especially in non-human primates. However, numerous human electroencephalography (EEG) and neuroimaging studies have shown both enhancement of primary and non-primary auditory cortex activation by selective auditory attention (Hillyard et al., 1973; Alho, 1992; Woldorff et al., 1993; Alcaini et al., 1994; Grady et al., 1997; Fujiwara et al., 1998; Jancke et al., 1999; Petkov et al., 2004; Rinne et al., 2007), and increased neural response selectivity/contrast (Ozaki et al., 2004; Ahveninen et al., 2006; Bidet-Caulet et al., 2007; Kauramaki et al., 2007; Okamoto et al., 2007; Neelon et al., 2011). Therefore,
focused auditory attention to a specific element (e.g. pitch) seems to augment neuronal selectivity for that element in the cortical region that it is preferentially processed in (for a review see (Fritz et al., 2007a). In support of this notion, single unit recordings in behaving animals have shown task-related bandwidth modulations or even “re-tuning” of receptive fields of primary auditory cortical neurons to behaviorally relevant frequencies (Edeline and Weinberger, 1993; Ohl and Scheich, 1996; Fritz et al., 2003; Fritz et al., 2005b; Fritz et al., 2007b; Atiani et al., 2009; Galindo-Leon et al., 2009). Furthermore, a recent non-human primate study demonstrated that enhancing and suppressive effects of attention are concurrently present, and that they are both “timed” to the predicted occurrence of attended stimuli. The central conclusion was that attention does not simply suppress responses to ignored stimuli, rather it serves to sharpen and stabilize the representation of attended stimuli, possibly by modulating the frequency tuning of neuronal ensembles (Lakatos et al., 2013a).

The sharpening of frequency tuning by “sideband” inhibition has long been observed, even in anesthetized or non-behaving animals (Shamma and Symmes, 1985; Suga, 1995; Sutter et al., 1999; Wang et al., 2000; Foeller et al., 2001; Sadagopan and Wang, 2010). It has been theorized that attentive listening may change the balance of excitatory and inhibitory inputs thereby augmenting sideband inhibition while also increasing response gain (Hromadka and Zador, 2007). An earlier study of population level sideband inhibition in passively behaving monkeys found that tones that resulted in the largest stimulus related inhibition, which were usually 2 octaves different in frequency from the tone which matched the best frequency (BF) of an A1 site, reset ongoing neuronal oscillations to their low excitability phases (i.e., the neurons in the local population were further from their firing threshold) (O’Connell et al., 2011). This process could serve as a dynamic mechanism during attention to augment feedforward sideband inhibition in A1, since it has been shown that the phase reset and
consequent entrainment of ongoing oscillatory neuronal activity is under strong attentional influence (Elhilali et al., 2009; Lakatos et al., 2009; Stefanics et al., 2010; Cravo et al., 2013; Lakatos et al., 2013a; Lakatos et al., 2013b). Therefore our overarching hypothesis is that these subthreshold modulatory mechanisms augment the fundamental stimulus related lemniscal feedforward excitation and inhibition related to attended stimuli to further sharpen the frequency tuning of A1, and enhance their sensory representation, which is supported by one of our recent studies mentioned above (Lakatos et al., 2013a).

The purpose of our current study was to parametrically test the extent of attention related effects on the activity of neuronal ensembles in A1. To achieve this, we presented the monkeys streams of pure tones in different blocks that differed in their frequency, covering the monkey’s entire hearing range (Pfingst et al., 1978). Our reasoning was that if we observe attention effects related to a large range of pure tones, this means that most of A1 is involved in the attentional modulation of sensory information, since frequency is projected onto cortical space in the auditory system (tonotopy). Alternatively, only the excitability of a relatively narrow A1 region would be modulated by attention to tones that corresponds to the BF and surrounding inhibitory sidebands (O’Connell et al., 2011; Lakatos et al., 2013a). Additionally, we were interested whether the responses of neuronal ensembles in different laminar locations, with differing connectivity patterns (feedforward vs. feedback), are differentially modulated by auditory attention. To answer these questions we analyzed laminar profiles of neuronal ensemble synaptic activity (indexed by current source density (CSD) analysis) and MUA recorded via linear array multielectrodes positioned in A1 of two macaque monkeys (*Macaca mulatta*) performing an intermodal selective attention task. We found that attending to auditory stimulus streams significantly sharpens frequency tuning, due to an increase in both response gain and selectivity. These effects are linked to the opposite sign predictive modulation of
neuronal excitability fluctuations (oscillations) in A1 sites “on” versus “off” the frequency representation for a given tone. Our results also indicate that while attention mostly boosts the neuronal responses to preferred stimuli in the infragranular and granular layers, attentional sharpening of frequency tuning is most apparent in the activity of supragranular neuronal ensembles.

3.3. RESULTS

In the present study, we analyzed laminar CSD and MUA profiles of responses to pure tones ranging from 353.5 Hz to 32 kHz in half octave frequency intervals (14 different frequencies) obtained with linear array multicontact electrodes. We recorded the neuroelectric activity of 39 primary auditory cortex sites which were distributed reasonably evenly along the tonotopic axis of A1, with best frequencies (BF) ranging from 0.5 kHz to 32 kHz. 15 (38.5%) of the sites had a BF <= 8 kHz while 24 (61.5%) of the sites had a BF > 8 kHz. During recordings, the two subjects had to perform a selective intermodal attention task: in separate blocks, the monkeys either had to attend to a rhythmic stream of auditory tone beeps and detect deviant tones that differed in their frequency (2-4 semitone difference from the standards) while ignoring visual stimuli (attend auditory or AA), or they had to attend to rhythmically presented LED light flashes and detect deviant flashes that differed from standards either in color or intensity (ignore auditory or IA) while ignoring stimuli in the auditory modality. The stimulus onset asynchrony (SOA) was kept constant in auditory and visual streams, and it differed slightly (auditory SOA = 624.5, visual SOA = 562.05) so that visual and auditory stimuli did not have a constant temporal relationship. This was meant to eliminate any multisensory “binding” effects, and facilitate the segregation of the two different modality streams. Repetition rates of stimuli in both auditory and visual modalities corresponded to the frequency of delta band oscillatory
activity (1.6 and 1.8 Hz respectively). The frequency of standard and deviant tones was the same within blocks, and was parametrically varied from block to block to sample neuronal activity related to both attended and ignored tones of all 14 different frequencies. The subjects were cued to which modality to attend to by presenting a singular stream in the “to be attended” modality.

Deviant or oddball stimuli in either modality’s stimulus stream occurred randomly at 3 to 6 second intervals; when occurring in a cued (attended) stream, these were targets. 80% of the targets were paired with a juice reward, while 20% were not. Similar to the behavioral paradigm used in one of our earlier studies (Lakatos et al., 2013a), the monkeys had to extend their tongues to obtain the juice as the spigot of the juicer was placed away from the monkey’s mouth. Licking on deviants both with and without paired reward was used to monitor performance (see Experimental Procedures).

3.3.1. Sharpening of Frequency Tuning by Attention

To ascertain whether the previously reported response gain (or enhancement) and suppression associated with attended preferred frequency and non-preferred frequency tones respectively (Lakatos et al., 2013a) was evident in our data, we first examined laminar CSD and MUA profiles in response to attended vs. ignored tones (Fig. 1). To facilitate the comparison of attended vs. ignored response amplitudes, we created frequency tuning curves by averaging event related laminar (averaged across all layers) MUA amplitudes in the 15-40ms post-stimulus timeframe (O’Connell et al., 2011), in response to the attended and ignored 14 different frequency tone streams. Figure 1 shows a representative example from a recording site with a BF of 8 kHz. The color maps show laminar profiles of CSD and MUA responses to attended (upper) and ignored (lower) BF and non-BF tones. We selected responses related to
non-BF tones with the largest suppressive attention effect (see tuning curves in Fig. 1). As expected based on the tuning curves, both CSD and MUA responses in the case of the BF tone streams were larger when the monkey was attending to the auditory modality. In contrast, but not surprisingly, since we selected responses to tones with largest attention related suppressive effect – the attended non-BF tone stream resulted in smaller CSD and MUA responses compared to when these same tones were ignored (same CSD and MUA scales). While the frequency difference between the tone resulting in the largest suppression effect and the BF varied from recording site to recording site (mean = 2.11 octave, standard deviation (SD) =

Figure 1  Representative laminar CSD and MUA profiles in response to attended and ignored BF and non-BF streams in A1. Schematic of a linear array multielectrode positioned in primary auditory cortex (A1). Cortical layers are indicated by numbers. To the right are current source density (CSD) and concomitant multiunit activity (MUA) response profiles related to the attended BF (8 kHz) and a non-BF (2 kHz) tone stream. Beneath are the CSD and MUA profiles evoked by the same tone streams while they were ignored. Note that there is a response enhancement for the BF tone in the attended vs. ignored condition, while in the case of the non-BF tone there appears to be a response suppression. Inset on left shows the frequency tuning curves created using laminar MUA amplitudes in response to 14 different frequency tones presented in blocks during both attentional conditions (attended-red, ignored-blue) for this particular A1 site. Arrows mark the frequencies of BF and non-BF tones. Note that while all tones with frequencies other than the BF can be considered non-BF, for this figure (and also Fig. 2a) we selected the non-BF tone related responses with the largest attention effect.
the largest response enhancement always occurred to attended BF tones, except for 3 sites (8% of all sites) where we did not observe an attention related response enhancement at the BF.

While we initially created frequency tuning curves using the 15-40 ms post stimulus time interval, since previous studies suggest that response to pure tones is largest in this “transient response” time-interval (Steinschneider et al., 2008; O'Connell et al., 2011), we next wanted to empirically test whether this was actually the timeframe in which the largest MUA attention effects arose. Figure 2A shows the pooled averaged laminar MUA responses to BF and non-BF tones (selected as described above) normalized to peak ignored BF response amplitude in each experiment, for all 39 experiments, in attend (red) vs. ignore auditory (blue) conditions. After statistically comparing MUA amplitudes recorded during the two attentional

![Figure 2](image.png)

**Figure 2** The effect of attention on the MUA response and sharpness of tuning. A) Pooled MUA responses to BF (top) and non-BF (middle) tone streams averaged across all layers during attend auditory (AA, red traces) and ignore auditory (IA, blue traces) conditions. The non-BF responses shown here were selected based on largest MUA response amplitude difference between attention conditions in each experiment. P-value graph (bottom) displays the result of AA vs. IA statistical comparison for each timepoint (Wilcoxon signed rank test). Even though the sign of the attention effect is different, the largest significant attention effect occurs around the same time from response onset (10 ms) to about 40 ms post-stimulus (marked with dotted green vertical lines) for both BF and non-BF responses. B) The frequency tuning curves display tuning curves pooled across all experiments in the AA and AI conditions. For each site (n=39) frequency tuning curves were created from MUA responses to streams of attended and ignored tones in the 10-40 ms post-stimulus timeframe. The tuning curves were then normalized to the value of the ignored BF MUA response measure, and shifted to align the BF of all sites (n=39) in the same position in the graph (BF). Stars indicate significantly different MUA amplitudes between the two conditions (Wilcoxon signed rank test). C) Boxplots show the pooled amplification indices (AA-IA BF related MUA response amplitudes - top) and suppression indices (Σ(AA-IA) all non-BF related MUA response amplitudes - bottom). The indices signal opposite sign attentional modulation (enhancement vs. suppression for BF vs. non-BF response related measures), both of which are significant (see results). D) Supragranular, granular and infragranular MUA responses associated with BF and selected non-BF stimuli averaged across all experiments. The 10-40 ms time interval is marked with dotted green vertical lines. E) Same modulation indices as in C) but separately for each layer.
conditions at each time point, it is clear that the largest significant attention effect occurs between the 10-40 ms time interval (marked with dotted green vertical lines) for both types of responses as shown by the p-value graph underneath (Wilcoxon signed rank test). Remarkably, significant differences in response amplitudes to attended vs. ignored BF and non-BF tones start before response onset, indicating that MUA is modulated predictively when anticipating BF vs. non-BF tones in A1.

Even though, as mentioned above, the largest attentional suppression effects occurred to tones on average two octaves different from each sites BF, we also wanted to characterize attention effects to the rest of the tones along the frequency tuning curve, as greater MUA suppression to attended tones with frequencies further from and closer to the BF than two octaves was evident in most experiments (e.g. Fig 1, tuning curves). Figure 2B shows the normalized pooled frequency tuning curve that was created from averaged laminar MUA amplitudes in the 10-40 ms timeframe. The frequency tuning curves of the 39 individual A1 sites were shifted to align the BF of all sites in the same position in the graph (BF), and all values were normalized to the ignored BF related laminar response amplitude in each experiment. Stars indicate significantly different attended vs. ignored MUA response amplitudes (Wilcoxon signed rank test, p < 0.01). It is apparent that in the case of BF tone streams, attending to the auditory modality resulted in a significant response amplitude increase (mean increase = 28.61%, SD = 38.61%), while for other tone streams significantly suppressive attention effects were detected across all experiments to tones with frequencies as close as 1 octave away from the BF, and as far away as 4.5 octaves. Note that at the extreme ends of the pooled tuning curves there were very few data points (large standard error in Fig. 2B), thus it is possible that suppressive effects extend even to frequencies further removed from the BF.
Results presented thus far indicate that selective attention to a stream of pure tones results in both the enhancement of BF tone related responses and a suppression of responses to most other frequency tones, and therefore, a sharpening of frequency tuning. To quantify these opposing effects, we devised two indices (Fig. 2C): 1) the amplification index is simply a subtraction of the ignored normalized (as above) BF-related MUA amplitude from the attended normalized BF related amplitude. If positive, this indicates an attention related enhancement of the response; this was observed in 36 (92.3%) of the experiments. 2) The suppression index is the subtraction of the sum of ignored non-BF tone related normalized MUA amplitudes from the sum of the attended non-BF tone related normalized MUA amplitudes. If negative, this index indicates a sharpening of tuning; we found this in 29 (74.36%) of the experiments. Thus, in most experiments, the net effect of attending to tones other than the BF is response suppression. As boxplots of the pooled indices show (Fig 2C), the amplification index (top) was significantly larger than zero (Wilcoxon signed rank test, $p < 0.001$) while the suppression index (bottom) was significantly smaller than zero (Wilcoxon signed rank test, $p = 0.0034$). The finding that in most A1 sites we found a positive amplification coupled with a negative suppression index indicates that selective auditory attention results in contrast gain (increased neuronal selectivity) as opposed to simply response gain (increased neuronal activity).

Thus far our analyses focused on MUA responses averaged across all cortical layers. To determine whether attention effects on MUA responses differed across layers, we selectively averaged MUA responses across supragranular, granular and infragranular electrode sites (see Experimental Procedures). Figure 2D displays these layer specific MUA responses related to BF and non-BF tones (selected the same way as in Fig. 2A). At first inspection it appears that in the case of the BF related responses the largest enhancement occurs in the middle and lower layers, while for non-BF related responses, suppression is most prevalent in the supragranular
layers. To better quantify layer specific attention effects, we calculated amplification and suppression indices for all layers separately (Fig. 2E). We found that as foreshadowed by the averaged MUA responses, while the amplification index was significant at a very conservative criterion for the granular and infragranular layers, it was significant only with a relaxed criterion for the supragranular layers (Wilcoxon signed rank test, gran: p<0.0001, infra: p<0.0001, supra: p=0.010). Additionally, the suppression index was only significantly smaller than zero for the supragranular layers (Wilcoxon signed rank test, supra: p=0.0009, gran: p=0.230, infra: p=0.204). This indicates that attention has differential effects on the representation of attended auditory stimuli in different layers; we will discuss this later.

3.3.2. Entrainment of Delta Oscillations by Attended Tone Streams
The purpose of our next set of analyses was to determine if the attention dependent entrainment of delta oscillations could be responsible for the attention related MUA effects seen above, as was suggested by a previous study (Lakatos et al., 2013a). If frequency specific delta entrainment is to serve as the mechanism of the predictive sharpening of tuning in A1, two basic predictions should hold true: 1) Supragranular delta oscillations should be entrained by most attended stimulus streams, and either the phase or strength of entrainment should differ for ignored streams. 2) Delta oscillations should be entrained to their high excitability phases by attended BF streams in order to predictively amplify responses, while they should be entrained to opposite phases for predictive suppression of non-BF stimulus streams.
To test these predictions, we examined the consistency and angle of supragranular delta phases in response to different frequency tone streams in the attended and ignored conditions. Figure 3A again shows MUA tuning curves from a representative site (BF = 4 kHz), with single trial supragranular delta (1.6 Hz corresponding to the repetition rate of the auditory streams) phases measured at stimulus onset for a subset of different frequency streams displayed as histograms below. The most apparent difference between delta phases related to attended and ignored streams is that, while in the attend auditory condition, delta phases are pooled around a certain phase value, phases appear completely random in the case when the same auditory

**Figure 3**  
Delta entrainment to different frequency tone streams.  

A) Overlaid frequency tuning curves, from a representative experiment, created using event related laminar MUA amplitudes (10-40ms), in response to AA (red) and IA (blue) tone streams. Bottom: Red histograms show the delta phase distribution across single trials at time of stimulus onset, related to a subset of AA tone streams (shown by arrows), black vertical lines show the mean phase. Blue histograms show the delta phase distribution across single trials related to the same tone streams but in the IA condition.  

B) Red trace shows the delta band inter-trial coherence (ITC) values for each attended tone stream in the same experiment. The purple dotted line denotes the average value above which ITC is significantly non-random. All 14 different frequency tones streams resulted in significantly biased delta phase distribution at this specific A1 site (100% significant delta ITC).  

C) To quantify the frequency dependent phase opposition shown in panel A), we subtracted the mean phase associated with the BF stream from all mean phases (this is why BF phase (green oval) is 0 in this graph). Next we determined how many of the non-BF tone streams resulted in a mean delta phase (red ovals) that was a half pi different from the BF phase (“outside” the half oscillatory cycle centered on the BF, marked by the blue dotted lines). In this specific case the measure was 86%.  


streams are ignored; this matches our first prediction. Additionally, an examination of the phase distributions related to each of the different frequency attended streams reveals that while in the case of the BF stream, phases are pooled between 0 and \(\pi\) [the downslope of the oscillation, associated with high excitability based on previous studies (Lakatos et al., 2005b; Lakatos et al., 2008; Lakatos et al., 2013a), for the non-BF tones phases are pooled oppositely, between \(-\pi\) and 0 (the upslope), satisfying our second prediction. This delta band phase opposition around the time of stimulus onset is also evident in the averaged CSD plots in Figure 1, as an opposite baseline fluctuation can be seen for attended stimuli: in the case of the BF stream a supragranular source over sink is visible just before 0, while in the case of the non-BF stream a sink over source is apparent. When the same stimuli are ignored this baseline fluctuation is absent, indicating a lack of entrainment in this condition.

To determine whether the above described delta phase effects hold true across all of our experiments, we first calculated the delta inter-trial coherence (ITC), quantifying phase-similarity across trials, for all attended and ignored streams in each experiment, and determined the ratio of significant ITC (Rayleigh’s uniformity test, \(p < 0.05\)) within each attention condition (Fig. 3B shows ITC values in the attend condition for same experiment as 3A). Next we calculated the mean phase of delta oscillations at stimulus onset for each stream, and determined the ratio of phase opposition compared to BF tone phase in each experiment amongst attended streams by subtracting the mean delta phase in response to the attended BF stream from the mean phase measured at stimulus onset for each of the attended non-BF tone streams. Graph in Figure 3C shows the results: in this experiment only one other tone stream resulted in a delta phase that fell within a half \(\pi\) (shown by the dotted blue lines) of the BF stream related phase (shown by green oval). In this specific case the “out of phase” measure was 86\%. 
Figure 4A shows these pooled delta phase related measures across all experiments. First, as predicted by previous studies (O’Connell et al., 2011; Lakatos et al., 2013a), delta phase at stimulus onset is significantly clustered around $\pi$ in the case of all attended BF tone streams (Rayleigh’s uniformity test, $p < 0.001$), thus indicating that attending to BF tone streams resulted in the entrainment of delta oscillations to their high excitability phase. Next, while irrespective of tone frequency, almost all attended tone streams resulted in significant delta ITC in all experiments indicative of entrainment (like in our illustrative experiment, Fig. 3B). Significant delta ITC was detected only in about 5% of the cases when tone streams were ignored (Fig. 4B). Finally, as Figure 4C shows that in nearly 80% of the cases, attended tone streams entrained delta oscillatory activity to a phase that was opposite to the BF stream related phase (mean = 77.29%, SD = 10.98%). To confirm that delta phase opposition is indeed related to the attentional sharpening of frequency, we split the experiments into three groups based on the ratio of phase opposition. 15 sites had smaller phase opposition ratio than the median (79%), 11 sites had larger, and the rest ($n = 13$) had median phase opposition ratios. Our prediction was that larger phase opposition ratios should result in larger suppression indices. 

![Figure 4](image_url)

**Figure 4  Pooled delta phase measures.** A) Pooled mean delta oscillatory phases of all sites associated with attended BF stimulus streams. B) Boxplots show % of tone streams across all experiments which resulted in significant delta ITC in the AA (red) and IA (blue) conditions. C) Boxplot shows % of attended tone streams in each experiment which entrained delta oscillations with a phase opposite to the BF stream related phase (calculated as described in Fig. 3C). D) Boxplots show the pooled suppression indices (see Fig. 2C) for A1 sites that showed greater, equal to or smaller phase opposition than the median phase opposition across all experiments (79%). The bracket indicates significant difference (Wilcoxon rank sum, $p<0.05$).
Figure 4D shows, this is exactly what we found: the suppression index for the “larger than median” group was significantly greater than the “smaller than median” group (Wilcoxon Rank Sum, p = 0.0127) indicating greater sharpening of tuning. The suppression index of A1 sites with median phase opposition was somewhere in-between (mean suppression indices for the smaller, equal to or larger than median phase opposition: -0.0627, -0.3717 & -0.6313). While this does not prove causality, it does however suggest that the frequency selectivity of delta entrainment is related to the sharpening of tuning in a given experiment.

Since human studies indicate that attended task structure related low frequency oscillatory activity is more prominent in the right hemisphere (Poeppel, 2003; Luo and Poeppel, 2012), and since a majority of our recordings were paired recordings in left and right A1 (14 paired recordings yielding 28 sites out of the total 39), we examined whether delta ITC related to attended streams was different across hemispheres. We found that delta ITC values averaged across different frequency attended streams for each recording site was virtually identical across hemispheres (left mean = 0.24, SD = 0.10; right mean 0.22, SD = 0.09), and was not significantly different (Wilcoxon rank sum, p = 0.11). Consequently, we expected that amplification and suppression indices should also be similar across hemispheres, which is exactly what we found (Wilcoxon rank sum, p = 0.33 and p = 0.29 respectively).

3.3.3. The Net Effect of Attention on the Excitability of A1 Neuronal Ensembles

The data described above show that, at least during a frequency discrimination task, delta oscillatory activity in an A1 site entrains with its low excitability phase to most rhythmic streams of tones covering the hearing range of the monkey, with the exception of tones that the given neuronal ensemble is tuned to. If true, this should be reflected in the modulation of baseline MUA and gamma frequency oscillatory activity, two measures that index the
excitability of a neuronal ensemble (e.g. Fries et al., 2001b; Lakatos et al., 2008). Specifically, while excitability should be predictively up-regulated in anticipation of BF tones, it should be down-modulated in anticipation of most other tones. This suggests a net predictive suppressive effect. To test this, we compared pre-stimulus activity associated with attended BF streams vs. averaged pre-stimulus activity associated with all tone streams (including the BF stream). Along with MUA and gamma oscillatory activity in the CSD, we also analyzed gamma in the local field potential (LFP), due to its implications for human electrocorticogram and scalp recordings.

Figure 5A shows non-baseline corrected laminar MUA, gamma CSD amplitude and gamma band LFP amplitude profiles from an A1 site related to BF tone streams (top), and related to all streams (averaged across all attended streams, including the BF stream) (bottom). It is apparent that while pre- and post-stimulus MUA is largest in the granular layer, gamma oscillatory activity has an additional supragranular maximum, similar to visual cortex (Buffalo et al., 2011; Spaak et al., 2012). This pattern was the most common, but varied considerably across recording sites. Since it appears that the excitability of all cortical layers tends to fluctuate largely synchronously (O’Connell et al., 2011; Lakatos et al., 2013a), and to reduce the complexity of the data, we averaged MUA and gamma amplitudes across all layers (Fig 5A, bottom).

In the MUA signal, there is a small fluctuation in baseline activity that is opposite in sign for BF and non-BF stimuli: for BF stimuli MUA was elevated immediately prior to stimulus onset, signaling enhanced excitability (Lakatos et al., 2005b; Lakatos et al., 2008) compared to the average activity related to all tone streams. This effect was even more obvious in the amplitude changes of gamma frequency range CSD and FP, which further supports the
The first column of colormaps shows MUA laminar profiles related to BF tone streams (top) and related to all tone streams (averaged), including the BF (bottom). Colormaps in the middle and to the right show gamma band activity (25 – 55 Hz) amplitude profiles extracted from CSD and local field potential (LFP) respectively. Traces on the bottom display the time course of MUA and gamma range activity averaged across all layers. Dotted vertical lines denote the immediate pre-stimulus (‘2’) and interstimulus (‘1’) timeframes used to calculate the modulation indices in B. Boxplots display the pooled difference of immediate pre-stimulus (-150 – -30 ms, marked by ‘2’) and interstimulus (-300– -150 ms, marked by ‘1’) for the three measures of neuronal activity. Inset p-values are the results of statistical analyses (Wilcoxon signed rank), testing whether the distributions are significantly different from zero. Note that in the case of BF stream related pre-stimulus activity, MUA and gamma are up-modulated towards the timing of attended stimuli, while in the net activity related to all streams these measures indexing the excitability of the local neuronal ensemble are down-modulated.
ensemble. To quantify this excitability modulation, we calculated a modulation index which is simply a subtraction of the interstimulus MUA or gamma band amplitude (marked as “1”, between green and black lines, -300 ms to -150 ms) from the immediate prestimulus amplitude (marked as “2”, between black and blue lines, -150 ms to -30 ms). If positive this index signals increasing MUA and gamma band activity in the prestimulus timeframe (predictive enhancement of excitability), while if negative it implies the reverse (predictive suppression of excitability). Boxplots in Figure 5B show the pooled modulation indices associated with BF and all tone streams for the three different neuronal measures. Statistical analyses comparing whether the distributions are significantly different from zero, which would signify no modulation, show that for all tone streams combined the modulation index is significantly smaller than zero for all three measures, while the MUA and gamma range FP modulation indices are significantly greater than zero for BF streams. These gamma band and MUA findings demonstrate two important effects: first, as suspected, there is a significant attention related, stimulus frequency-dependent difference in pre-stimulus excitability which aids in suppressing or enhancing later stimulus-related responses. Second, as predicted by results of recent findings in humans (Lakatos et al., 2013b), when a subject attends to a stream of pure tones the net effect across the tonotopic surface of A1 is predictive suppression, as the vast majority of A1 neurons are off-BF for any given tone.

3.4. DISCUSSION

The present study found that during a selective auditory attention task, frequency tuning was sharpened all along the tonotopic axis of A1. Attention both amplified excitatory responses to pure tone streams whose frequency matched a site’s BF and increased response suppression for tone streams whose frequency differed from the BF by 1 to 4.5 octaves. At all recording sites,
amplification and suppression were accompanied by an entrainment of low frequency oscillatory activity to opposing, high vs. low excitability phases when the monkey attended to BF vs. non-BF tone streams, respectively. We found that entrainment occurred to almost all attended tone streams, and that the majority of streams entrained delta oscillations to their low excitability phases. Due to a projection of frequency onto cortical space in A1, these findings provide indirect evidence that most of A1 entrains to its low excitability phase when pure tones are attended. In support of this notion, our results also demonstrate that the net effect of attention on the excitability of neuronal ensembles tuned to different frequencies across A1 is suppression. Additionally we found opposing superficial to lower layer gradients in the suppressive vs. enhancement effects of attention: while the degree of response enhancement related to attended BF tones increased from supra- towards infragranular layers, suppression increased in the opposite direction. Taken together our findings provide convincing evidence that when auditory stimuli are presented in rhythmic streams, attention sharpens the frequency representation of task-relevant items using the predictive modulation of cortical excitability, and this impacts neuronal responses in a layer specific way.

3.4.1. Mechanism of Increased Response Selectivity in A1

While our study does not show causality, the close correspondence between response enhancement vs. suppression and the opposing phases at which delta oscillations entrain suggests a strong connection between these phenomena. If we take into account the prior findings that the phase of lower frequency ongoing and entrained neuronal oscillations modulates the amplitude of higher frequency oscillations and MUA, both of which reflect excitability changes of the local neuronal ensemble (Fries et al., 2001b; Lakatos et al., 2005b; Lakatos et al., 2008; Whittingstall and Logothetis, 2009), the most likely mechanistic
relationship between the attention related effects in our study is the following: attended tone streams entrain ongoing oscillations to their high excitability phases in regions that preferentially process the attended frequency content, while ongoing low excitability fluctuations of neuronal ensembles outside this “BF region” are entrained to their low excitability phases, minimizing the effect of auditory inputs in non-attended frequency channels (not tested here but see Lakatos et al., 2013a). This mechanism would be especially useful in noisy environments, where it would act as a narrow spectrotemporal bandpass filter. There are at least two distinct ways this mechanism could emerge when attending to a stream of tones: either predictive enhancement and suppression via opposite phase delta entrainment is “set up” simultaneously by modulating inputs mediating counterphase phase reset and thereby the excitability of neuronal ensembles independently, or more likely the predictive enhancement of a given frequency channel results in suppression of all others. Whichever the case, there are also at least three anatomical routes through which attention can result in frequency dependent entrainment: top-down modulation of phase reset via cortico-cortical feedback connections, horizontal modulation of phase reset via either excitatory or inhibitory connections (this would only work if suppression is a “consequence” of enhancement) or modulation of non-specific thalamocortical inputs either via intrathalamic connections or corticothalamic feedback. Disentangling which of these functional-anatomical routes is the main culprit in orchestrating the predictive spatiotemporal modulation of differently tuned neuronal ensembles in A1 will likely require a combination of techniques such as electrophysiological recordings merged with electrical microstimulation, optogenetics, pharmacological manipulations and computational modeling.
3.4.2. Layer Specificity of Attention Effects

Our study took an important first step in trying to unravel the functional circuitry underlying the mechanism of selective auditory attention in A1 by demonstrating layer specific attentional effects: specifically we found that MUA suppression is largest in the supragranular layers while enhancement dominated in the granular and infragranular layers. It is known that in both the visual and auditory cortices the supragranular layers receive a large amount of input from local and long range horizontal connections (Gilbert and Wiesel, 1983;Ts'0 et al., 1986;Ojima et al., 1991;Wallace et al., 1991;Bosking et al., 1997;Ojima and Takayanagi, 2004). Accordingly recent studies conducted in the primary auditory cortex of both anesthetized and awake passive mice have shown degraded frequency selectivity in the supragranular layer compared to other layers (Guo et al., 2012;Winkowski and Kanold, 2013), which is thought to be due to projections from other spectrally distinct columns in A1 (Kaur et al., 2004;Kaur et al., 2005;Happel et al., 2010;Moeller et al., 2010). This exact connectivity to disparately tuned neuronal ensembles in the supragranular layers might enable attention to orchestrate opposite phase effects that are temporally linked across A1. In support of this a recent study showed that modification of supragranular layer responses in animals trained on a tone detection task vs. naïve animals was due to intracortical connections originating from a spectral distance of 1 octave or more (Guo et al., 2013). It has also been shown that layer I contains the highest density of GABAergic (inhibitory) synaptic endings, and that the source of these could be layer I horizontal cells which have long lateral axonal branches that radiate within layer I (Prieto et al., 1994). In addition, more recent studies have demonstrated the existence of long range horizontal intracortical inhibitory connections in layer II (Tomioka et al., 2005), and long range (up to 3 mm) excitatory projections terminating on inhibitory interneurons especially in the
supragranular layers (Kurt et al., 2008). Either of these suppressive pathways could explain our results.

Aside from horizontal cortico-cortical connections, an alternative possibility is that these opposite phase effects are orchestrated by non-specific thalamocortical inputs which are known to target the supragranular layer (Molinari et al., 1995; Jones, 1998; Huang and Winer, 2000). Whatever the mechanism, since the supragranular layers of A1 project overwhelmingly to higher order cortical regions, the attention related response selectivity in these layers will result in sharpened feedforward output from the primary auditory cortex.

In contrast, our results indicate that the main attention effect in the granular and infragranular layers is response enhancement. The infragranular layers of A1 are known to be part of the corticofugal system and its corticothalamic projections feedback to all divisions of the medial geniculate body (MGB), while it has fewer projections targeting the inferior colliculus (Winer and Prieto, 2001; Winer, 2005). Therefore it is possible that the attentional response gain we observe in the BF region’s infragranular layer serves mainly to increase the responsiveness and selectivity of the collicular and thalamic regions processing the ascending sensory input, and indeed the presence of topographically organized corticothalamic fibers (Winer et al., 2001) and frequency-specific corticothalamic modulation (Zhang and Suga, 2000) supports this. For example it has been shown that electrical stimulation of layer V neurons increased the threshold and sharpened the frequency tuning curve of recorded inferior colliculus neurons (Sun et al., 1996). Also inactivation of A1 decreased the spontaneous activity and modified the bandwidths of most single units in all subdivisions of the MGB, suggesting that as well as having an excitatory effect, corticofugal projections may support adaptive filtering at the thalamic level by regulating cell’s response properties by changing their bandwidth responsiveness (Villa et al., 1991; Zhang, 1997; Zhang and Yan, 2008). Another
possibility is that corticothalamic feedback projections are boosted at the attended frequency channel to support some sort of a “winner take all” mechanism via collaterals to the reticular nucleus of the thalamus, or via the corticothalamic projections themselves onto inhibitory neurons of the MGB.

3.4.3. The Effect of Attention on Auditory Stimulus Processing

Previous studies from our group (O’Connell et al., 2011; Lakatos et al., 2013a) have shown that non-BF tones approximately two octaves different in frequency from the BF tone reset or entrain ongoing oscillations to the low excitability phase. The current study extends that finding by demonstrating that attention amplifies these suppressive effects, and by confirming that the majority of A1 entrains with its low excitability phase to any non-preferred tone stream. Nonetheless, the largest attentional response suppression occurs to tone streams about two octaves away from the BF, just as it does in passive monkeys (O’Connell et al., 2011), and the response enhancement “bandwidth” around the BF was about 1 octave. This is in general agreement with the finding of Atiani and colleagues (Atiani et al., 2009) who found response enhancement in cells whose BF was within 0.6 octaves (bandwidth ~1.2 octaves) of the target tone during an easy tone detection task and suppression when the target tone’s frequency was >0.6 octaves in A1 of behaving ferrets relative to their passive state. Likewise a recent study in rat A1 showed that attention always enhanced cells responses to their BF, and only enhanced target responses if the target frequency was within 1 octave of the cells BF (Jaramillo and Zador, 2011). Importantly, our results go further to suggest that the excitatory bandwidth varies across cortical layers: upper layers appear to have a narrower bandwidth of attentional enhancement, that in the lower layers appears broader.
Previous studies investigating receptive field (RF) alterations using conditioning paradigms mainly focusing on the changes in responses to target stimuli (e.g. CS+) have reported quite a few different types of RF modulation, sometimes conflicting. There is, however, a general agreement that attention serves to increase the contrast/salience of behaviorally relevant stimuli (Edeline and Weinberger, 1993; Ohl and Scheich, 1996; Fritz et al., 2003; Fritz et al., 2005b). A previous study from our lab (Lakatos et al., 2013a) and the results of the present study support this consensus. An important difference is that previous studies used single unit firing as an index of frequency tuning changes, our study examined neuronal ensemble activity using CSD analysis, which is a more sensitive indicator of small changes in response amplitude and subthreshold excitability changes. Our results indicate that the spectrotemporal filter mechanism (Lakatos et al., 2013a) is not restricted to the vicinity of neuronal ensembles processing the attended frequency content, rather, it extends to all of A1, as proposed by a recent human study (Lakatos et al., 2013b). Since most of A1 consists of non-BF neuronal ensembles in the case of narrowband frequency content, the net effect of attention on subthreshold neuronal activity is suppressive (see Fig. 5). It will be up to future studies to determine whether and how changes in the bandwidth of the frequency content of attended stimuli change the shape of the filter mechanism. In support of flexible “filter shape”, results from Fritz et al. provide evidence for multiple passbands (Fritz et al., 2007c).

Besides animal studies, recent human neuroimaging studies using masking paradigms have also provided evidence that both response gain and attenuation work in concert during auditory attention to sharpen the frequency tuning of auditory cortex (Kauramaki et al., 2007; Okamoto et al., 2007; Okamoto et al., 2009b; Ahveninen et al., 2011). While we cannot compare the results of these types of paradigms directly with our study, we predict that adding
notched noise to pure tones would even further increase our effects, since a tonic drive to non-BF sites would make the subthreshold suppression effects more robust.

Our results can also explain the somewhat counterintuitive finding that in scalp EEG studies using constant predictable auditory SOAs attention does not result in an amplitude increase of the auditory event related potential (Schwartze et al., 2013;Lakatos et al., 2013b), specifically the N1 which is thought to originate mostly in supratemporal auditory cortex (Liegeois-Chauvel et al., 1994;Godey et al., 2001). An examination of the amplification and suppression indices in Fig. 2C reveals that they are very similar in amplitude. Indeed, while the amplification index signals an approximately 28% response increase related to BF tones, the suppression index reveals a 33% response suppression summed across all other tones. Since scalp recorded activity reflects the sum of responses across all of auditory cortex, these two effects likely cancel each other, resulting in no response amplitude change. On the other hand, a plethora of human scalp EEG studies have shown a general enhancement of the N1 component during auditory attention. This contradiction may be explained in most cases by the fact that since many of these studies used random SOAs in their stimulus presentation (Woldorf et al., 1993;Fujiwara et al., 1998;Kauramaki et al., 2007;Okamoto et al., 2007;Okamoto et al., 2009b;Ahveninen et al., 2011), low frequency neuronal oscillations in auditory cortex may not have been able to entrain and predictably modulate the excitability of disparately-tuned neuronal ensembles. Although a recent human fMRI study does provide convincing evidence that frequency selective response enhancement and suppression operates in non-rhythmic auditory tasks as well (Da Costa, 2013), this continuous processing mode (Schroeder and Lakatos, 2009) might not be as efficient in suppressing responses related to non-relevant stimuli, since in this case the temporal dimension of the spectrotemporal filter mechanism cannot operate (Lakatos et al., 2013a). It is worth mentioning that the neural generators of the
scalp ERP components have been localized to primary and non-primary auditory cortices and it is not known if oscillatory excitability is modulated by attention the same way in non-primary auditory areas as it is in A1. We speculate that “subthreshold spectral filters” likely do not play an important role in higher order auditory regions, since the frequency specificity of neuronal ensembles in these regions is poor (Rauschecker, 1995; Rauschecker et al., 1997; Kaas and Hackett, 2000; Rauschecker, 2004; Lakatos et al., 2005a).

3.4.4. Context and Content

In speech, the processing of information content depends on the temporal context (Ahissar et al., 2001; Ghitza, 2011). Our results add to the mounting evidence that similarly, the processing of specific sensory inputs relaying information about the physical properties of the external world are modulated by the internal neurophysiological context. This context appears to consist of rhythmic excitability fluctuations on multiple timescales that are coupled to each other in space and time (Buzsaki and Chrobak, 1995; Buzsaki and Draguhn, 2004; Lakatos et al., 2005b; Buzsaki, 2010). Our present experiments nicely illustrate the difference between content and context: while content is transmitted through inputs that match the physical properties of the stimulus (in the case of pure tones through thalamocortical inputs tuned to the frequency of these), the context is modulated across a broader range of neuronal ensembles in order to enhance valuable and suppress interfering information. The way this is achieved by the brain is that it models or maps the fundamental properties (timing and frequency) of auditory stimuli onto cortical space in the form of subthreshold excitability fluctuations. This internal map or context is tied to the timing of the external temporal context which makes it effective in adaptively filtering relevant information. It is easy to see how a failure to form an accurate internal representation via a “subthreshold excitability phase map” or a failure of the alignment
mechanism between internal and external context would lead to less effective stimulus processing, and such failures likely contribute to cognitive deficits in schizophrenia patients (Lakatos et al., 2013b).

3.4.5. Conclusions

Our results provide evidence that auditory attention results in frequency and cortical layer specific response gain and increased response selectivity. It appears that the mechanism is a predictive modulation of cortical excitability via the entrainment of neuronal oscillations. Frequency dependent entrainment enhances excitability in BF and suppresses excitability in non-BF tuned neuronal ensembles across all of A1 at time-points when attended stimuli are predicted to occur (context), thereby enabling predictive stimulus specific modulation of driving type inputs (content) and leading to sharper frequency tuning. In cases when the bandwidth of the attended frequency content is narrow (pure tones), the net effect of selective auditory attention is suppression. Our finding of layer specific attentional effects suggests that while functionally boosting the firing of neuronal ensembles processing relevant stimulus properties might be more important for feedback, filtering out irrelevant information may be more important for feedforward information transfer.

3.5. EXPERIMENTAL PROCEDURES

3.5.1. Subjects. In the present study, we analysed the electrophysiological data recorded during 39 penetrations of area A1 of the auditory cortex of 2 female rhesus macaques (19 and 20 penetrations) weighing 5-7 kg, who had been prepared surgically for chronic awake electrophysiological recordings. Prior to surgery, each animal was adapted to a custom fitted
primate chair and to the recording chamber. All procedures were approved in advance by the Animal Care and Use Committee of the Nathan Kline Institute.

3.5.2. Surgery. Preparation of subjects for chronic awake intracortical recording was performed using aseptic techniques, under general anesthesia, as described previously (Schroeder et al., 1998). The tissue overlying the calvarium was resected and appropriate portions of the cranium were removed. The neocortex and overlying dura were left intact. To provide access to the brain and to promote an orderly pattern of sampling across the surface of the auditory areas, plastic recording chambers (Crist Instruments) were positioned normal to the cortical surface of the superior temporal plane for orthogonal penetration of area A1, as determined by pre-implant MRI. Together with socketed Plexiglas bars (to permit painless head restraint), they were secured to the skull with orthopedic screws and embedded in dental acrylic. A recovery time of six weeks was allowed before we began data collection.

3.5.3. Electrophysiology. Animals sat in a primate chair in a dark, isolated, electrically shielded, sound-attenuated chamber with head fixed in position, and were monitored with infrared cameras. Neuroelectric activity was obtained using linear array multi-contact electrodes (23 contacts, 100 µm intercontact spacing). These multielectrodes were inserted acutely through guide tube grid inserts, lowered through the dura into the brain, and positioned such that the electrode channels would span all layers of the cortex (Fig. 1), which was determined by inspecting the laminar response profile to binaural broadband noise bursts. In 14 experiments (7 in each subject) we recorded the neuronal activity of primary auditory cortex in the left and right hemispheres simultaneously. Neuroelectric signals were impedance matched with a pre-amplifier (10x gain, bandpass dc-10 kHz) situated on the electrode, and after further amplification (500x) they were recorded continuously with a 0.01 - 8000 Hz bandpass digitized with a sampling rate of 20 kHz and precision of 16-bits using custom made software in
Labview. The signal was split into the field potential (0.1-300Hz) and MUA (300-5000Hz) range by zero phase shift digital filtering. MUA data was also rectified in order to improve the estimation of firing of the local neuronal ensemble (Legatt et al., 1980). One-dimensional current source density (CSD) profiles were calculated from the local field potential profiles using a three-point formula for the calculation of the second spatial derivative of voltage (Freeman and Nicholson, 1975). The advantage of CSD profiles is that they are not affected by volume conduction like the local field potentials, and they also provide a more direct index of the location, direction, and density of the net transmembrane current flow (Mitzdorf, 1985; Schroeder et al., 1998). At the beginning of each experimental session, after refining the electrode position in the neocortex, we established the best frequency (BF) of the recording site using a “suprathreshold” method (Steinschneider et al., 1995; Lakatos et al., 2005a). The method entails presentation of a stimulus train consisting of 100 random order occurrences of a broadband noise burst and pure tone stimuli with frequencies ranging from 353.5 Hz to 32 kHz in half octave steps (duration: 100 ms, r/f time: 5 ms, SOA = 624.5). Auditory stimuli were produced using Tucker Davis Technology’s System III coupled with MF-1 free field speakers.

3.5.4. Behavioral task and stimuli. We trained 2 monkeys to perform an intermodal selective attention oddball task, which required them to attend to one modality, and discriminate stimuli within that same modality, while ignoring stimuli in the other modality. In this paradigm, auditory and visual stimulus streams were presented simultaneously, and monkeys were either cued to detect frequency deviants occurring at random time intervals in the auditory stream, or a color difference in rhythmically flashing LEDs in the visual stream. The auditory stream consisted of pure tone beeps at 40 dB SPL (25 ms duration, 5 ms rise/fall time) with a constant stimulus onset asynchrony (SOA) of 624.5ms (1.6Hz). The SOA was set so that the repetition rate would correspond to the delta frequency range of ongoing neuronal activity. The frequency
of the auditory standards was parametrically varied across blocks in half octave steps between 0.3-32 kHz, so we had 14 different frequency tone streams. Frequency deviants occurred in the stream of standard tones every 3-9 seconds randomly. To get the monkeys to attend to the rhythmic streams of tones, in the beginning of training, 0.25-1 ml juice reward was delivered to them simultaneously with each deviant through a spout. The spout was positioned such that the monkeys had to stick out their tongue in order to get the juice. Licking was monitored using a simple contact detector circuit (Slotnick, 2009), the output of which was continuously recorded with Labview together with the timing of standard and deviant tones for offline analyses. In this phase of training the frequency difference between the standard and deviant tones was about one octave. After 2 sessions, the juice reward was omitted on every 10th deviant. If the monkeys licked on these deviants without a paired juice reward, signalling that they were attending to the tones and actively discriminating the deviants, we omitted the reward on 20% of the deviants, and also gradually lowered the frequency difference to 2-4 semitones. After 10-20 sessions on average, the monkey’s performance became relatively stable: they were reliably licking on juiceless deviants before the next stimulus occurred in the train. When subjects became satiated, they stopped licking even when juice was delivered; this usually occurred after more than 500 deviants.

In the attend visual condition, we presented a rhythmically flashing green LED with an SOA of 562.05ms (1.8Hz). In this case the deviant was a change in color of the LED to red (in the beginning of training/experiments) or a more intense green flash which occurred every 3-9 seconds randomly. To increase task difficulty we decreased the intensity difference between standard and deviant green flashes. The subjects were cued to attend to one of the streams by the preceding cueing stream that matched the modality of the stream to be attended to. The subjects always responded to deviants in only one of the streams, never to deviants in both
streams. One of the subjects performed this task 76% correct, while the other monkey only 64% correct, as determined by licking on juiceless deviants, which remained stable throughout the course of all experiments. There was no difference between behavioural performances in the auditory vs. the visual task. We only analysed standards that preceded deviants (with or without juice) on which the subjects licked.

3.5.5. Data analysis. Data were analysed offline using native and custom-written functions in Matlab (Mathworks, Natick, MA). After selective averaging of the CSD and MUA responses to the tones presented in the suprathreshold tonotopy paradigm, recording sites were functionally defined as belonging to AI or belt auditory cortices based on the sharpness of frequency tuning, the inspection of the tonotopic progression across adjacent sites, and relative sensitivity to pure tones versus broad-band noise of equivalent intensity (Merzenich and Brugge, 1973; Rauschecker et al., 1997; Lakatos et al., 2005a). In the present study only recordings obtained from area A1 were analysed. At the end of each animal’s experimental participation, functional assignment of the recording sites was confirmed histologically (Schroeder et al., 2001).

All analyses were conducted on the neural responses to standard stimuli and the responses to the first three standards after each deviant were excluded due to motion artifacts (licking) and due to the fact that deviant stimuli could never occur in these stimulus positions. Utilizing the BF-tone related laminar CSD profile, the functional identification of the supragranular, granular and infragranular cortical layers in area A1 (see Fig. 1) is straightforward based on our earlier studies (Schroeder et al., 1998; Schroeder et al., 2001; Lakatos et al., 2005a). In the present study we focused the analyses of ongoing and event related neuronal activity on the supragranular CSD with largest BF tone related activation (sink), and initially the MUA averaged across all layers. The reason for this selection is that
both ongoing and entrained oscillatory activity are most prominent in the supragranular layer 
(Lakatos et al., 2005b; Lakatos et al., 2007; Lakatos et al., 2008; Lakatos et al., 2009), and they 
appear to reflect synchronous excitability fluctuations of the local neuronal ensembles across 
all layers, as evidenced by synchronous MUA amplitude fluctuation across the layers 
(O'Connell et al., 2011; Lakatos et al., 2013a). Also, dominant delta frequency neuronal activity 
in all cortical layers is largely coherent with supragranular delta oscillatory activity (O'Connell 
et al., 2011; Lakatos et al., 2013a), albeit actual phase values signaling high or low excitability 
are different at different laminar locations (see Lakatos et al., 2013a Supplementary Fig. 1). For 
the analysis of laminar response amplitude effects, we averaged MUA activity across 
electrodes spanning the supragranular, granular and infragranular layers (on average 8.26, 3.72 
and 5.2 electrodes respectively).

To be able to determine the phase relationship of delta oscillatory activity to the timing 
of attended and ignored stimuli in stimulus streams, instantaneous phase in single trials was 
extracted by wavelet decomposition (Morlet wavelet) on 135 scales from 0.5 to 3.2 Hz (see e.g. 
Fig. 3B). Independent of their frequency composition, cyclically occurring events like the 
suprathreshold, “evoked type” response waveforms can artificially bias phase measures at the 
frequency that corresponds to the stimulus presentation rate (see Lakatos et al., 2013a 
Supplementary Fig. 3 for examples and further explanation). To minimize this bias, a linear 
interpolation was applied to the single trials prior to wavelet analysis in the 5 – 150 ms time 
interval which in the case of most BF tones contained evoked-type activation (Lakatos et al., 
2013a). To characterize delta phase distribution across trials, the wavelet transformed single 
trial data was normalized (unit vectors), the trials were averaged, and the length (modulus) of 
the resulting vector was computed (Lakatos et al., 2007). The value of the mean resultant 
length, also called inter-trial coherence (ITC) ranges from 0 to 1; higher values indicate that the
observations (oscillatory phase at a given time-point across trials) are clustered more closely around the mean than lower values (phase distribution is biased). Phase distributions were evaluated statistically using circular statistical methods. Significant deviation from uniform (random) phase distribution was tested with Rayleigh’s uniformity test. The $\alpha$ value was set at 0.01 for all statistical tests.

Gamma amplitudes (Fig. 5) were extracted from CSD and LFP signals by first band-pass filtering in the 25-55 Hz band, and then calculating the analytic amplitude of the signal using the Hilbert transform in each single trial before averaging.
CHAPTER 3

Multi-scale entrainment of coupled neuronal oscillations in primary auditory cortex.

O’Connell M.N.1,2, Barczak A.1,4, Ross D.1, Schroeder C.E.1,2,3, Lakatos P.1,3

1Cognitive Neuroscience and Schizophrenia Program, Nathan Kline Institute, Orangeburg, New York 10962, 2 Program in Cognitive Neuroscience, City College CUNY, New York, New York 10031, 3Department of Psychiatry, Columbia College of Physicians and Surgeons, New York, New York 10032, 4Department of Neuroscience & Physiology, New York University School of Medicine, New York NY
4.1. SUMMARY

Earlier studies demonstrate that when the frequency of rhythmically-presented pure-tone stimuli is task relevant, the ongoing excitability fluctuations (oscillations) of neuronal ensembles in primary auditory cortex (A1) entrain to stimulation in a way that sharpens frequency tuning. This study examined how neuronal activity is modulated if only the temporal features of broadband stimulus streams are relevant. We presented macaques with auditory clicks arranged in 33 Hz (gamma timescale) streams, repeated at a 1.6 Hz (delta timescale) rate. Such multi-scale, hierarchically organized temporal structure is characteristic of vocalizations and other natural stimuli. Monkeys were required to detect and respond to random deviations in the temporal pattern of gamma rate clicks. As expected, engagement in the auditory task resulted in the multi-scale entrainment of delta- and gamma-band neuronal oscillations across all of A1. Surprisingly, however, the phase-alignment, and thus, the physiological impact of entrainment differentiated across the tonotopic map in A1. In the region of 11-16 kHz representation, entrainment most often aligned high excitability oscillation phases with the task-relevant events in the input stream. In the remainder of the A1 sites, entrainment generally resulted in response suppression. Our data indicate that the suppressive effects were due to the low excitability phases that delta oscillations entrained to and the phase amplitude coupling of delta and gamma oscillations. Regardless of the phase or frequency, entrainment appeared stronger in left A1, indicative of the hemispheric lateralization of auditory function.
4.2. INTRODUCTION

The most fundamental organizing principle of the auditory system at lower hierarchical levels of acoustic information processing is a faithful spatial representation of the auditory receptor surface in the cochlea (Schreiner and Winer, 2007). One of the likely reasons for the topographical organization of auditory information (tonotopy) across several earlier processing stages is that just like in signal processing (e.g. EEG analysis, photo or music editing), information can be best manipulated (e.g. filtered or sharpened) at high resolutions to enhance relevant features. Once the information is compressed, only larger facets can be manipulated. Since frequency representation is condensed to a large degree already at the level of the second cortical processing stage in belt auditory cortex, theoretically, any refinement in the frequency composition of the auditory environment should ideally take place before this stage, in primary auditory cortex (A1) or subcortical structures.

Several studies have found that along with amplifying responses to task-relevant stimulus frequencies, attention also suppresses responses in neuronal ensembles tuned to the “ignored region” of the frequency spectrum (Fritz et al., 2003;Fritz et al., 2005a;Fritz et al., 2007c;Da Costa et al., 2013;Lakatos et al., 2013a). This essentially represents a spectral filter mechanism that sharpens the frequency tuning of A1 by modulating auditory information across topographically organized neuronal ensembles. Two of our recent studies have found that when band limited attended auditory stimuli (pure tones) are presented rhythmically, a temporal filter component is superimposed on this spectral filter, in that attended frequency content will only be amplified and frequency tuning sharpened at specific times, when relevant stimuli are predicted to occur (Chapter 2;Lakatos et al., 2013a). The mechanism of this spectrotemporal filter is the entrainment of ongoing neuronal oscillatory activity that represents spontaneous excitability fluctuations of the local neuronal ensemble. Oscillatory entrainment
by the attended stimuli results in a predictive excitability modulation across all of A1. A key to the mechanism of the spectral filter component is that neuronal oscillations are entrained in counterphase across A1 neuronal ensembles tuned to relevant vs. irrelevant frequency content: while the excitability of neuronal ensembles tuned to attended frequencies is up-regulated preceding the predicted occurrence of stimuli to amplify responses to relevant frequency content, the excitability of neuronal ensembles around this region is down-regulated, in order to suppress irrelevant inputs that temporally coincide with attended stimuli.

Contrasting with this counterphase entrainment in A1, an earlier study investigating the processing of rhythmic visual stimuli found that in the neuronal ensembles of the primary visual cortex (V1), ongoing oscillations were always entrained to their high excitability phases by attended stimuli (Lakatos et al., 2008). A likely reason for the difference between entrainment effects across topographically organized neuronal ensembles in A1 and V1 is that while the auditory studies used pure tones which excite only a subset of the tonotopically organized neuronal ensembles in A1, the visual study used flashes that were not confined in space, and thus activated a large proportion of the retinotopically organized V1 neuronal ensembles. Also, while in the auditory tasks the topographically mapped feature, the frequency of the stimuli, was task relevant, in the visual task, the topographically mapped feature, the spatial location of the flash was task irrelevant.

Based on these earlier results, the main hypothesis we tested in the present study was that if subjects attend to broadband auditory stimuli, whose frequency content is task irrelevant, ongoing oscillations in most auditory neuronal ensembles will be entrained to their high excitability, depolarizing phases, in order to predictively amplify incoming inputs. We also hypothesized that as a consequence, the overall effect of attention to these stimuli will be a response enhancement across A1 sites independent of tonotopy. To test this, we presented
auditory click trains, since clicks have a broad frequency spectrum. The same five clicks formed both frequently occurring standard, and rarely occurring deviant, or target stimuli, and the only difference between these was in their temporal structure, rendering the frequency content task irrelevant. Contrary to our hypothesis, effects of engagement in the auditory task differentiated across the tonotopic gradient in A1. Only neuronal ensembles tuned to around 11-16 kHz had a strong tendency to entrain to their high excitability phases on multiple timescales. The more common effect, observed over the remainder of A1 sites, was response attenuation, due to the predictive suppression of neuronal excitability on at least one of the task structure related timescales (intra and inter click train) in most A1 sites. Additionally, we found evidence that entrainment was left lateralized, indicating that similar to humans (for a review see (Giraud and Poeppel, 2012), auditory cortical function might be lateralized even at the level of primary auditory cortex in non-human primates.

4.3. RESULTS

We analyzed neuroelectric data recorded in 48 A1 sites of 2 macaque monkeys. Ongoing and event related neuronal activity was recorded with linear array multielectrodes which spanned all cortical layers at each A1 recording site. To be able to directly compare simultaneous activity of left and right hemisphere A1 neuronal ensembles, the majority of the data (42 sites in 21 experiments) was obtained via simultaneous left and right A1 recordings targeting regions tuned to similar frequencies. To minimize the effects of volume conduction and more precisely define local laminar transmembrane current flow profiles (Freeman and Nicholson, 1975; Mitzdorf, 1985; Schroeder et al., 1998), we calculated one dimensional current source density (CSD) from the field potentials and carried out most of our analyses on the CSD waveforms and concomitant multiunit activity (MUA).
Auditory stimulus-related neuronal activity was recorded in two conditions in separate trial blocks: either the monkeys were attending to frequently repeating standard stimuli in order to detect deviants that differed from standards in their temporal structure (engaged), or they were passively listening to the same stimuli (passive). The stimulus onset asynchrony (SOA) was a constant 624.5 ms in both conditions, corresponding to the average wavelength of dominant delta frequency oscillations in the ongoing neuronal activity of primary auditory cortex (Lakatos et al., 2005b). Standard auditory stimuli used in the experiments consisted of 5 clicks arranged at regular, 30.3 ms time intervals (corresponding to 33 Hz), while deviant stimuli (targets in the engaged condition) differed in that the 3rd click was shifted towards the 4th (shift = 15–30.3 ms, Fig. 1). The 33 Hz repetition rate of the 5 clicks corresponds to the gamma frequency range of the EEG. This stimulus arrangement resulted in a hierarchically organized rhythmic stimulus structure on two coupled time scales (i.e. delta (1.6 Hz) and gamma (33 Hz), that was designed to examine whether the entrainment of ongoing neuronal oscillations can occur simultaneously in multiple

**Figure 1. The auditory stimulus stream.** A) We presented auditory clicks that had a hierarchically organized rhythmic structure on two time-scales: trains of 5 clicks had a repetition rate of 33 Hz, while the click trains occurred at a rate of 1.6 Hz. Deviant (target in the engaged condition) stimuli had a different temporal structure: the 3rd click was shifted in time. B) To the left, a standard click train recorded with a microphone and digitized at 100 kHz. To the right, the waveform of an individual “click” as emitted by the speakers. The “wavelength” of the bipolar pulse is approximately 0.3 ms.
frequency bands, a mechanism that was proposed as one of the cornerstones of speech perception and analysis (Schroeder et al., 2008; Ghitza, 2011; Giraud and Poeppel, 2012).

4.3.1. The Effect of Engagement on Responses to Click Trains

To assess the general effect of engagement in the task on auditory responses, we statistically compared MUA response amplitudes to the gamma quintets in engaged vs. passive conditions within each experiment. Since we presented stimuli at two different loudness levels in both conditions (40dB and 50dB), we determined the effect of engagement separately for these. Across all recording sites, response onset latency to 50 dB attended click trains was on average 7.41 ms (standard deviation (SD) = 0.92 ms) and did not differ significantly between the active and passive conditions (Wilcoxon signed rank, p = 0.239). Since previous studies have shown that response onset varies across differently tuned regions in A1 (Mendelson et al., 1997; Kaur et al., 2004; Lakatos et al., 2005a; O'Connell et al., 2011), we tested this by comparing response onsets in A1 regions with best frequencies (BF) of 8 kHz or lower to response onsets in neuronal ensembles tuned to higher frequencies (BF > 8 kHz). As predicted, response onset latencies in A1 sites tuned to lower frequencies were significantly longer than those in sites tuned to higher frequencies (17 vs. 31 recording sites, Wilcoxon rank sum, p = 0.0409). Interestingly, when we tested whether response onset latency was significantly different across left and right hemispheres (23 vs. 25 recording sites), we found that left hemisphere latencies were significantly shorter, albeit only on average by 0.74 ms (Wilcoxon rank sum, p = 0.0227).
Since, as described above, there was no significant difference in response onsets across different behavioral conditions, we measured MUA response amplitude averaged across all cortical layers in the timeframe from earliest response onset (6 ms) until 40 ms post-stimulus (160 ms). When we statistically compared response amplitudes within all experiments in the engaged vs. passive conditions (Wilcoxon rank sum), we found that for both 40dB and 50dB click trains, engagement resulted in significant response suppression in most A1 sites ($n = 26$ (54%) in the case of 40 dB and $n = 32$ (66%) in the case of 50 dB click trains). The upper traces in Figure 2A (upper panel) show the averaged responses of sites that showed significant engagement related response suppression to the stimulus trains presented at either loudness ($n = 32$, “suppression group”). The rest of the sites either showed no engagement-related response modulation at any loudness ($n = 12$), or a significant response enhancement ($n = 3$ in the case of 40dB and $n = 2$ in the case of 50dB trains). Since enhancement only occurred in a fraction of our

**Figure 2. Engagement in a task generally suppresses responses to click trains.**

A) Multiunit responses to trains of 5 auditory clicks (33 Hz repetition rate). Top traces show averaged responses to click trains presented at 40 and 50 dB loudness (dark vs. light blue), in engaged and passive trial blocks (solid vs. dotted traces), recorded in A1 sites where engagement resulted in a significant suppression of the MUA response in the 6 – 160 ms time range. The bottom traces are pooled responses from A1 sites where engagement resulted in either no amplitude change (12 out of 16) or a significant enhancement of MUA in the same time range. B) The best frequency (BF) of sites with significant engagement related suppression (blue) and no significant suppression (violet). Note that the latter group of sites is concentrated in the high frequency tuned region of A1.
experiments (0.08%), we pooled these sites with the “no response amplitude change” ones and termed them “no suppression group” (n = 16), the averaged responses of which are shown in the lower panel of Figure 2A. By observing the averaged responses of the suppression group (Fig. 2A upper), we noted that while the effect of loudness only appears to affect the transient parts of the response, the effect of engagement is observable across the whole “response timeframe”, even in the periods between transient responses to clicks. We also noted that in these averaged responses, the effect of engagement on the transient responses to clicks corresponds to a 10 dB decrease in loudness. A third relevant observation is that, whereas response amplitudes are larger to the first and second clicks, they reach a “steady state” following the 3rd click in the train.

To determine whether there is a relationship between the tuning of the neuronal ensembles and effect of engagement on click train related responses, we sorted the recording sites according to their best frequency, which is displayed in Figure 2B. It is apparent that the “no suppression” group of sites mostly had BFs of 11 or 16 kHz, and never BFs lower than 5.6 kHz. The distribution of the sites with engagement related suppression appears more evenly distributed along the tonotopic axis. The two groups of sites were relatively evenly distributed across both hemispheres (10 vs. 6 non-suppressive sites in left vs. right hemispheres).

When we examined the responses in laminar CSD profiles in an attempt to categorize them based on the effect of engagement, similar to the MUA, we at first did not notice any apparent pattern, in fact we were puzzled by the great variety of engagement related effects. Figure 3 illustrates this by showing the CSD response profiles of 3 differently tuned A1 sites: one from the non-suppressive and two from the suppressive group. The CSD profiles of the non-suppressive site (BF = 16 kHz) are highly similar in the passive vs. engaged conditions
(Fig. 3A), with slightly higher averaged response amplitudes in the supragranular layers, as illustrated by the traces to the right that show the CSD of selected electrodes from different cortical layers. Additionally, more noticeable in the CSD traces is that the baseline appears “more tilted” in the engaged condition in the supragranular layers. The MUA response of this

Figure 3. Representative laminar response profiles. A) Color-maps to the left display laminar CSD profiles in response to click trains from an A1 site with a BF of 16 kHz (no suppression group). The boundaries of the supragranular, granular and infragranular layers are marked by black horizontal lines. The click train is represented by black lines on the x-axis. Note that there appear to be sink(red)-source(blue) pairs representing active and associated passive currents in each of the layers. Color-maps to the right display the concomitant laminar MUA responses, which show a somewhat larger amplitude response in the engaged condition. The upper 3 traces to the right of the color-maps show CSD responses from the supragranular, granular and infragranular electrode locations marked by numbers to the left of the CSD maps (for the selection of these see methods). The apparently largest response amplitude difference in passive vs. engaged condition is at the supragranular electrode location. Note that at stimulus onset, the baseline is negative-trending (especially in the engaged condition), and at the onset of the fourth click (orange line at 90.9 ms) the gamma frequency waveform (SSR) in between is also negative-trending. The bottom traces display the MUA averaged across all layers recorded in passive vs. engaged conditions. B) Same as A), but from a low BF A1 site with significant engagement related suppression (suppression group). As opposed to A), the supragranular CSD in the engaged condition is positive trending at 0 ms, indicating an opposite phase low frequency excitability modulation. The slope of the SSR at 90 ms is negative trending. C) Same as A), but from a relatively high BF A1 site with significant engagement related suppression. While similar to A), the baseline is negative trending, the SSR waveform is positive trending at the onset of the fourth click. This latter effect appears much stronger in the engaged condition. Note that in all three sites, MUA at stimulus onset is oppositely trending to the supragranular CSD, indicating that similar to what previous studies found, a negative CSD trend signals increased, while a positive CSD trend signals decreased excitability.
particular site is enhanced in the engaged condition across all cortical layers. The next site (Fig. 3B) is tuned to low frequencies (BF = 0.5 kHz), and as the laminar MUA profiles show, the MUA response to click trains is suppressed across all layers in the engaged condition. Compared to the first site, the laminar CSD response in the passive condition appears overall larger in amplitude, with maybe a slight polarity difference in the supragranular layers. Similar to the first site, the baseline appears more tilted on the selected supragranular channel in the engaged condition (CSD traces), although in the opposite direction. The third site shown in Figure 3 (Fig. 3C) was tuned to 8 kHz, and since this site also belongs to the suppressive group, the MUA response appears attenuated across all layers in the engaged condition. The most apparent difference between the laminar CSD response profiles in the two conditions is that in the supragranular layers, the source over sink pattern in the passive condition appears flipped to sink over source in the engaged condition.

4.3.2. The Pattern of Delta and Gamma Frequency Entrainment across A1

To quantify the observed CSD differences between the two conditions, we measured the mean phase and phase consistency (inter-trial phase coherence or ITC) of neuronal activity at the delta and gamma frequencies that corresponded to the repetition rates across and within click trains (1.6 and 33 Hz respectively). Our reasoning was that several previous studies have shown that modulating the phase and/or strength of oscillatory entrainment can modulate responses to attended tones (Lakatos et al., 2008; Lakatos et al., 2013a; Chapter 2). Thus, assuming that the phases measured reflect the phase of entrained oscillatory activity as opposed to evoked type, de novo generated neuronal activity, the pattern of phase alignments could reveal the mechanism of response suppression in the engaged condition. To verify this assumption, we compared the amplitudes of delta and gamma band neuronal activity in data that were recorded
in the absence of stimulation (spontaneous activity) to delta and gamma amplitudes measured in the passive and engaged conditions.

Figure 4A shows the spectrograms of supragranular neuronal activity (CSD) in the absence of stimulation and in different task conditions during the presentation of click trains. While it is obvious that at both delta and gamma stimulation rates, the amplitude spectrum of neuronal activity is “peaked” compared to the spontaneous spectrum, note that this is paired with lower amplitudes around the peak in the auditory stimulus stream related spectra resulting in no significant net amplitude change in the delta and gamma bands (Kruskal-Wallis test, $p = 0.9519$ and $p = 0.1549$ respectively, Fig. 4B). Rather, the peaks represent a reorganization of oscillatory activity to match relevant temporal scales, that results in a concentration of energy at the frequencies that correspond to the repetition rates of stimuli. In other words, the peaks mostly signal neuronal activity that is less variable in frequency in the delta and gamma bands, which has been shown to be characteristic of oscillatory entrainment (Lakatos et al., 2013a; Zoefel and Heil, 2013). In fact Figure 3B of Lakatos 2013a displays...
spectrograms of spontaneous neuronal activity following auditory stimulus presentation, with
delta peaks still corresponding to the rate of stimulus stream presentation evident, which are
very similar to delta peaks in the stimulus-related spectrograms of our Figure 4A. Nevertheless,
we cannot exclude the possibility that evoked type activity contributes to the measured spectra.
As a matter of fact it is likely, especially in the case of 50 dB click trains: the harmonic at
double the stimulation rate (~ 3.2 Hz) can be a strong indication of evoked type activity that
“distorts” the sinusoidal waveform that is characteristic of entrainment. Previous studies
indicate that the amplitude ratio of evoked type (added) neuronal activity to the ongoing
(entrained) neuronal activity determines the “distorting” effect of evoked responses on phase
measurements of the ongoing neuronal oscillations (Lakatos et al., 2013a). Since based on the
spectra, this ratio is very small in the case of 40 dB click trains, we analyzed delta and gamma
phases related to these lower intensity stimuli in engaged vs. passive conditions. To avoid
confounding effects of the larger evoked response to the click train, as in previous studies
(Lakatos et al., 2013a; Chapter 2), we applied linear interpolation to the data in the 5 – 150 ms
timeframe before we measured delta phases (see Experimental procedures). Furthermore, while
we measured delta phase at stimulus onset (0 ms), gamma phases were measured at the time of
the fourth click (90.9 ms) to get a more reliable estimate of the entrained (steady state) gamma
phase.

Figure 5A displays the histograms of mean delta (top) and gamma (bottom) phases
across all experiments in the engaged (left) and passive (right) task conditions. It is apparent
that the phase distributions are bimodal in most cases (except gamma phases in the passive task
condition). One group of phases peaks between –pi and 0, on the upward deflection of the
neuronal oscillation, while the other group is centered on the downward deflection. Our
previous studies analyzing supragranular CSD oscillations in the same laminar position have
Figure 5. The distribution of pre-stimulus delta and gamma oscillatory phases across differently tuned A1 regions. A) Histograms show the distribution of delta and gamma phases measured at the onset of the first and fourth click (0 and 90.9 ms respectively) in the engaged and passive conditions. Phases were measured at the delta and gamma frequencies that correspond to the SOA between and repetition rate within click trains (1.6 and 33 Hz respectively). Black traces above the histograms represent one oscillatory cycle. It is apparent that both delta and gamma phases in the engaged, and delta phases in the passive condition have a bimodal distribution. While half of the phase is pooled on the positive-trending (hyperpolarizing, blue), the other half is pooled around the negative-trending (depolarizing, red) phase. B) The BF distribution of sites with hyperpolarizing and depolarizing pre-stimulus phases. Note that sites with depolarizing pre-stimulus delta phases in the engaged condition have high BFs. Also note that the gamma phases in high frequency tuned sites are mixed, while low frequency sites have depolarizing gamma phases at the onset of the fourth click. Importantly, the BF distribution of sites with depolarizing and hyperpolarizing pre-stimulus phases is markedly different. C) Theoretically, delta and gamma phases can combine in four different ways, and as the histograms show, they indeed do. Note that the distribution of sites with depolarizing delta and gamma phases (delta>0 & gamma>0) in the engaged condition appears to be very similar to the BF distribution of sites with no significant engagement related suppression (Figure 2B).
spontaneous (incidental) neuronal ensemble firing and gamma oscillatory amplitudes, both of which are highest on the depolarizing phase of ongoing oscillations. To verify this in our data, we first binned MUA and gamma oscillatory amplitude (measured in the 25 -50 Hz band) averaged across all layers into two bins based on the phase of delta oscillatory activity: MUA and gamma during delta phases from −π to 0 (upward deflection) fell into one bin, while the rest (during delta phases from 0 to −π, the downward deflection) were put into the second bin. We found that even though the difference was on average very small (0.8% difference for MUA and 2.9% for gamma), both MUA and gamma frequency laminar activity was significantly larger during the downward slope of the supragranular delta oscillation (Wilcoxon signed rank, both p < 0.0001), confirming that this is indeed the high excitability or depolarizing phase of delta. Similarly although to a lesser degree, we found that MUA was significantly higher in amplitude during the depolarizing phase of gamma band oscillatory activity (Wilcoxon signed rank, p = 0.015).

To examine the relationship between the frequency tuning of A1 neuronal ensembles and the phase of delta and gamma entrainment, we created bar graphs with the phase of oscillations color coded (Fig. 5B, red = depolarizing, blue = hyperpolarizing phase). We found that in the engaged condition, attended click trains in sites with higher BFs entrained delta oscillations to their depolarizing phase, while in sites tuned to lower frequencies delta oscillations were entrained to their hyperpolarizing, low excitability phase. The mean phase of gamma oscillatory activity showed a more complicated pattern: in sites tuned to ≤ 2 kHz and to 11 -16 kHz we measured depolarizing phases, while hyperpolarizing gamma phases appeared biased towards sites tuned to frequencies surrounding 11 – 16 kHz. In the passive condition, mean phase distributions across differently tuned sites appeared less congregated around specific frequencies.
Next we examined how the phases of delta and gamma entrainment “combine” in each site. One possibility is that, for example, the depolarizing phase of delta always co-occurs with the depolarizing phase of gamma in one group of sites, and the hyperpolarizing phases of the entrained oscillations combine in another group of sites. However, two other combinations are theoretically possible: hyperpolarizing gamma phases combined with depolarizing delta, and vice versa. In fact, when we looked at the combination of delta and gamma phases, we found that all four possible combinations occur. Furthermore these seem to be grouped in sites tuned to similar frequencies (Fig. 5C): e.g. while hyperpolarizing delta and depolarizing gamma phases co-occur in regions tuned to 2 kHz and below, depolarizing delta and gamma phases co-occur overwhelmingly in regions tuned to 11-16 kHz. We noted that the BF distribution of sites in this latter group shows remarkable similarity to the BF distribution of sites in the no-suppression group (Fig. 2B). When we compared the “depolarizing delta-gamma” (n = 15) and no-suppression group of sites (n = 16), we found that 12 of the sites were indeed the same. This indicates that the majority of sites in the depolarizing group show either no response amplitude change or a response enhancement in the engaged vs. passive condition. It also follows that sites in the other three phase combination groups belong to the group of sites with significant engagement related MUA response suppression. To verify this and to uncover any multiscale entrainment specific differences, we pooled MUA responses based on delta-gamma phase combination into four groups, and compared response amplitudes in the steady state portion of the response (60 – 160 ms) in the engaged vs. passive conditions (Fig. 6). We also used the 60-160 ms timeframe to avoid responses to the first two clicks, since gamma entrainment is most likely only developed after the second or third click. We found that as predicted, with the exception of the depolarizing delta-gamma group, engagement resulted in significant response suppression.
Figure 6. MUA responses in A1 regions with differing pre-stimulus phases. Pooled MUA responses of sites with different pre-stimulus delta-gamma phase combinations. Boxplots to the right show the MUA response amplitudes across sites averaged in the 60 – 160 ms time interval. While the first group of sites shows no engagement related suppression, the other three groups of sites do. Note that the patterns of suppression seem to match the delta-gamma phase combinations remarkably well: while the suppression is maximal at the times of the individual click related responses in the depolarizing delta – hyperpolarizing gamma group (upper right panel, light blue), the suppression is more sustained in the groups where pre-stimulus delta is in hyperpolarizing phase (lower panels, cyan and dark blue). The group with hyperpolarizing delta-gamma phase combination (cyan) shows the most significant response suppression.

A comparison of the phase combinations in the differently tuned sites (Fig. 5C) reveals that the mechanism is partly an engagement related phase flip from depolarizing (passive condition) to hyperpolarizing (engaged condition) both in the case of delta and gamma oscillations (e.g. some sites tuned to 0.7 kHz, 1.4 kHz and 4 kHz). An opposite engagement related phase flip can be observed in some regions tuned 11 – 16 kHz. Nonetheless, the mean delta and gamma phase in the majority of sites does not change (e.g. sites tuned to 0.5 kHz), thus an engagement related change in the phase of entrainment can only partially explain the observed effects on MUA responses. Thus we asked whether phase consistency across single trials was different in engaged vs. passive conditions, since it has been shown that engagement results in a stronger phase reset and entrainment of ongoing neuronal activity (Lakatos et al., 2009, 2013a), and a stronger enforcement of suppressive phase patterns via entrainment could in theory result in significant suppression of responses. We found that as expected, both delta (1.6 Hz) ITC measured at stimulus onset, and gamma (33 Hz) ITC measured at the time of the
4th click in the click train was significantly greater in the engaged condition (Wilcoxon signed rank, both p < 0.001). While in the passive condition, delta and gamma ITC was not always significant (Rayleigh test, p < 0.05 in 29 sites for delta and 38 sites for gamma), delta and gamma phase consistency was significant in all sites in the engaged condition.

4.3.3. Engagement Related Hemispheric Asymmetry

Recent human research suggests that the processing of auditory stimuli structured at different timescales is hemispherically asymmetric (for a review see Giraud and Poeppel, 2012). We designed our stimuli in part to mimic the multi-temporal scale organization aspect of vocalizations, and thus, we were interested in the question of whether there is evidence of hemispheric asymmetry in the entrainment of fast and slow oscillations. To determine this, we

![Figure 7. Hemispheric lateralization of intertrial phase coherence and the amplitude of delta and gamma oscillations. A) Pooled delta and gamma (1.6 and 33 Hz) ITC values in left and right A1 sites in engaged and passive conditions. Generally, engagement results in a significant ITC increase (see results), and both delta and gamma ITC is higher in left hemisphere sites. Brackets mark significant differences (Tukey's test, p < 0.01). B) Pooled delta and gamma oscillatory amplitudes in left and right hemisphere A1 recordings in engaged and passive conditions. Both delta and gamma amplitude values follow the same trend as ITC, but none of the effects are significant. C) Spectrograms, of ongoing and 40 dB click train stream related activity in the engaged condition.](image)
pooled our delta and gamma ITC measures according to task condition and hemisphere. As Figure 7A shows, we found that delta ITC related to click train streams in the engaged condition was significantly greater in left hemisphere sites than delta ITC in either left or right hemisphere sites in the passive condition (Tukey’s test, $p < 0.01$). Importantly, left delta ITC in the engaged condition was also greater than right delta ITC in the same condition. These data indicate that there is a hemispheric asymmetry in the strength of delta entrainment due to engagement. A similar trend is apparent for gamma ITC, however left hemisphere gamma ITC in the engaged condition was only significantly larger than right hemisphere gamma ITC in the passive condition (Tukey’s test, $p < 0.01$). When we compared oscillatory amplitudes across hemispheres and task conditions, at the delta and gamma frequencies that corresponded to the SOA (1.6 Hz) and the repetition rate of click in the gamma quintets (33 Hz), there was a trend towards a similar leftward bias but no significant effects (Kruskal-Wallis test, $p > 0.05$) (Fig. 7B). Taken together these data show that left A1 exhibits greater stimulus structure related delta and gamma band oscillatory activity, and that engagement in the task enhances hemispheric differences in the oscillatory neuronal activity of the supragranular layers.

Important, as our results above foreshadowed (Fig. 4), larger delta and gamma amplitudes related to click trains in the left A1 are not due to larger evoked responses. As the spontaneous and stimulus-related spectra in Figure 7C show, the difference between spontaneous and auditory stimulus-related delta amplitudes at the rate of stimulation is actually larger in the right hemisphere indicating that perhaps in the right hemisphere evoked activity contributes more substantially to the delta peak in the spectrum of stimulus-related activity. Contrary to this, in the left hemisphere there is no net amplitude change between the two conditions, indicating that most likely entrainment is responsible for the delta peak at the stimulation rate. Additionally, there appears to be a hemispheric asymmetry in the spectra of
spontaneous neuronal oscillations, with larger amplitude supragranular activity in the left hemisphere overall which may contribute to the larger stimulus-related oscillatory amplitudes in the delta and gamma bands seen in Fig. 7B. This effect is especially evident in the delta frequency range (1 – 2.5 Hz), yet the amplitude difference between hemispheres was not significant (Wilcoxon rank sum, p = 0.0823).

4.4. DISCUSSION

Our main hypothesis was that when the broadband frequency spectrum of attended stimuli is irrelevant for an auditory task, ongoing oscillatory activity across all of A1 would be entrained to its high excitability, depolarizing phase, maximally amplifying auditory responses. However, to our surprise we found that even though all of A1 entrained its ongoing neuronal oscillations to the temporal structure of attended stimuli on two timescales, delta and gamma, the net effect of entrainment on auditory responses was mostly suppressive. Both the pattern of entrainment and engagement related response amplitude modulation differentiated across the tonotopic map in A1: ongoing oscillations of most neuronal ensembles in the 11-16 kHz region of A1 were entrained to their high excitability phases when monkeys engaged in the auditory task, and responses to task relevant stimuli in these sites were either enhanced or not significantly modulated compared to responses in the passive condition. In contrast, in most neuronal ensembles outside this A1 region, either delta (sites further from the 11-16 kHz region) or gamma (sites closer to the 11-16 kHz region) oscillations were entrained to their low excitability phases in the engaged condition (see Fig. 5B), which co-occurred with significant response suppression compared to the passive condition. Congruent with a more organized pattern of entrained delta and gamma phases, stimulus timing related delta and gamma phase consistency (ITC) were both significantly larger in the engaged compared to the passive
condition. Taken together, our findings indicate that neuronal ensembles tuned to the higher part of the audible spectrum play a central role in the sensory representation and processing of relevant broadband transient sounds, like auditory clicks. Additionally we found that engagement-related oscillatory entrainment on both time scales was stronger in left hemisphere A1 sites.

4.4.1. Mechanisms of Predictive Response Suppression

So how does the multiscale entrainment of delta and gamma band oscillations result in a net suppressive effect in most neuronal ensembles? Previous studies provide a wealth of evidence, which we verified in the present study, that both delta and gamma oscillations have depolarizing (or high-excitability) and hyperpolarizing (or low excitability) phases (for a review see Young and Eggermont, 2009). Our results show that in about half of the A1 sites examined, delta and gamma band oscillations were entrained to the former, while the other half to the latter (Fig. 5A). On a first hunch, this should result in an equal distribution of response enhancement and suppression across sites, which is not what we found (16/32 ratio of no-suppression vs. suppression sites). The reason for this is twofold: first, hyperpolarizing and depolarizing phases of entrained delta and gamma oscillations are not always paired; we found that they co-occur in all four possible combinations. Second, delta and gamma oscillations are phase amplitude coupled, meaning that the phase (i.e. high/low excitability) of a lower frequency oscillation determines the amplitude (large/small) of a higher frequency band oscillation (Lakatos et al., 2005b; Lakatos et al., 2008), as shown by the significantly smaller gamma amplitude on the hyperpolarizing phases of delta oscillations in our data. In light of the delta-gamma phase-amplitude coupling, let’s consider the four phase combinations. If the depolarizing phases of delta and gamma co-occur, this should result in high excitability at times
when stimuli are predicted to occur, due to a summation of depolarizing effects (e.g. Fig 6, upper left panel). However, if gamma oscillations entrained to their hyperpolarizing phases ride on the depolarizing phase of delta, gamma will be large amplitude. Thus, the hyperpolarizing phase of gamma will negate the depolarizing effect of delta, resulting in a net hyperpolarized state of the local neuronal ensemble in short, precisely timed temporal windows of low excitability centered on the clicks (e.g. Fig 6, upper right panel). In the remaining two categories of sites, since delta is entrained to its hyperpolarizing phase at which gamma amplitudes are low, the phase of gamma does not play an effective role in modulating excitability. Therefore, the net effect will be long time-scale predictive suppression related to the hyperpolarizing phase of delta. In support of this, while mainly the transient part of the click train related responses is suppressed in sites with depolarizing delta and hyperpolarizing gamma entrainment (Fig 6, upper right panel), suppressive effects appear much more tonic (longer time-constant) in hyperpolarizing delta sites (Fig 6, lower two panels).

4.4.2 The Effect of Engagement on Neuronal Activity in A1

Contrasting task engaged and passive conditions, like in the present study, is often used in animal studies to investigate behavioral state related changes in neuronal activity. Regardless of sensory modality, a common finding in rodent studies when comparing responses to stimuli in engaged vs. passive states is that responses are suppressed in the active behavioral condition (Fanselow and Nicolelis, 1999; Castro-Alamancos, 2004; Crochet and Petersen, 2006; Ferezou et al., 2006; Otazu et al., 2009). This is usually interpreted as a sharpening or refinement of the sensory input. Our main finding is in line with these previous studies in that the overall effect of engagement in the task is response suppression. However, contrary to results of a previous study in rat auditory cortex (Otazu et al., 2009), this suppression appears to be tonotopically
organized across A1 neuronal ensembles: it occurred mostly in A1 neuronal ensembles that were tuned to frequencies higher or lower than 11-16 kHz. Our data also reveal a candidate mechanism for the engagement related sharpening of the sensory representation: the modulation of subthreshold neuronal ensemble activity via the alignment of rhythmic excitability fluctuations to the temporal structure of relevant auditory stimuli. This alignment, the entrainment of neuronal oscillations occurs in the passive condition as well (similar to Lakatos et al., 2005b), but to a significantly lesser degree, and in a less organized pattern.

4.4.3. The Importance of Broadband Transient Sounds in Auditory Processing

In one of our earlier studies we proposed that the brain “models” the spectrotemporal properties of selectively attended auditory stimuli and stimulus streams in the form of temporally evolving phase patterns across topographically organized A1 neuronal ensembles (Lakatos et al., 2013a). This in turn forms the basis for enhancing and stabilizing the representation of attended auditory information at the expense of irrelevant, background auditory stimuli. The present study however found that the physical frequency spectrum of the auditory click is represented in a “distorted” form, since its representation is mostly enhanced in high, while suppressed in low frequency regions of A1. Thus we speculate that it is possible that sharp transients, like clicks or formant transitions represent a special category of stimuli, for which preserving an accurate frequency representation is less important. Rather, the main role of these “acoustic edges” could be to orchestrate the coherent multiscale entrainment of neuronal oscillations across differently tuned A1 neuronal ensembles, thereby setting up a spatiotemporal excitability pattern that is ideal for the parsing and processing of relevant auditory content mainly contained in the lower frequency spectrum e.g. speech (Fletcher, 1948;Peelle et al., 2013). In
this theoretical framework, acoustic edges would form the temporal context that enables the most efficient processing of the acoustic content by modulating ongoing neuronal oscillations.

Broadband transient sounds are common features of speech in humans (e.g. broadband consonants) and conspecific vocalizations in monkeys (May et al., 1989; Wang et al., 1995). Aside from communication sounds, they also occur frequently in the acoustic environment, in which case they mostly indicate something alerting requiring quick action (e.g. the snap of a twig). Thus, in theory, it would be advantageous to process these sounds via a fast dedicated auditory processing hierarchy of neuronal ensembles. Indeed, there are neurons in the posteroventral cochlear nucleus that are specialized in responding to broadband transients, called octopus cells. The main function of the octopus cells appears to be the integration of the cochlear activation via the summation of orderly dendritic synaptic activation, which compensates for the travelling wave delay of the cochlea (Rhode et al., 1983; Golding et al., 1999; Oertel et al., 2000; McGinley et al., 2012). These cells fire extremely fast and temporally very precisely (Rhode and Smith, 1986). Their output is transmitted via a separate ascending pathway mainly to the contralateral ventral nucleus of the lateral lemniscus, a pathway which appears to be much more prominent in humans (Adams, 1997). Interestingly, it has also been shown that octopus cells integrate cochlear inputs over about 1/3 of the audible spectrum (Oertel et al., 1990; Golding et al., 1995; Golding et al., 1999), which does correspond to the BF “spread” of the no-suppression group in our data. Thus, we hypothesize that the group of non-suppressive sites in A1 that are tuned to 11-16 kHz might form the first cortical stage of the ascending auditory pathway specialized in rapidly processing broadband transient sounds.

Besides rapid alerting, this “transient specialized system” together with the modulation of ongoing neuronal activity across A1 could play a crucial role in the processing of complex acoustic patterns like communication sounds. A central role could be parsing (Ghitza and
Greenberg, 2009; Buzsaki, 2010; Ghitza, 2011; Giraud and Poeppel, 2012): entrained delta and gamma phase related suppression of the neuronal ensembles processing speech sounds (<5 kHz) would be an efficient mechanism to segment continuous speech on two time scales, corresponding to the length of phonemes and syllables. Since speech has regularly and predictably interchanging broadband transients (i.e. consonants) and more “tonal” elements whose main spectral energy is band limited and is usually below 5 kHz (i.e. vowels), besides parsing, the other main role of transients in speech might be to prepare lower frequency cortical areas for the processing of band-limited sounds. Our results provide evidence that this could be achieved by resetting and entraining ongoing oscillatory activity to their low excitability phases in most A1 regions: as a consequence, the high excitability, depolarizing phase of ongoing oscillations will be centered on acoustic elements positioned between sharp transients (acoustic edges). Based on our previous studies (Chapter 2; Lakatos et al., 2013a), lower frequency speech elements (vowels) could also reset delta and gamma oscillations to their depolarizing phases in low and to their opposite, hyperpolarizing phases in high frequency regions, which in turn would prepare these for an upcoming high frequency element, or sharp transient. Thus we hypothesize that during speech processing, counterphase entrained oscillations would be reset twice during an oscillatory cycle, once by high frequency and one by low frequency inputs. This would provide a highly adaptive dual timing mechanism for the synchronization of neuronal oscillations to attended speech that is thought to be a key element of speech processing and perception (see below). This mechanism should be especially helpful in noisy environments, like at a cocktail party. An everyday observation in support of this is that it is close to impossible to make out someone’s speech on a phone (which transmits acoustic signals only below 5 kHz) when background noise is high or when multiple people are speaking. We speculate that this is most likely due to the missing temporal context contained in the high
frequencies of the auditory spectrum. A similar mechanism could explain aging related deficits in speech comprehension, which manifests stronger in the presence of environmental noise, since ageing often results in high frequency hearing loss (reviewed by Pichora-Fuller and Souza, 2003).

4.4.4. Evidence for Hemispheric Functional Lateralization in Non-Human Primates

Both delta and gamma oscillations, along with theta have been proposed to be important in the processing of speech and species specific communication (Schroeder et al., 2008; Ghitza, 2011; Giraud and Poeppel, 2012). We found that multiscale oscillatory entrainment at these rates shows greater phase consistency, at the time attended auditory stimuli occur, in left A1, indicating a stronger involvement of left hemisphere oscillatory activity. This finding provides support for the functional asymmetry of left and right auditory systems at the level of their first cortical processing stage, primary auditory cortex, possibly indicating that the precursor of left hemisphere association with speech is present in non-human primates. Previous studies in monkeys provide behavioral (Petersen et al., 1978; Ghazanfar et al., 2001), ablation (Heffner and Heffner, 1984) and neuroimaging (Poremba et al., 2004) evidence for hemispheric lateralization for the processing of species specific communication. Anatomical studies in new and old world monkeys also found evidence for a leftward asymmetry (Heilbroner and Holloway, 1988; Gannon et al., 1998; Gannon et al., 2008). Nevertheless, since the results relating to functional asymmetry are scarce, the hypothesized functional lateralization of auditory function is still unresolved in non-human primates. To our knowledge, our study is the first to provide electrophysiological evidence for hemispheric lateralization in monkeys. Importantly, our left and right measures are directly comparable, since most data was recorded simultaneously in left and right primary auditory cortices.
We found that the hemispheric asymmetry in the strength of delta and gamma entrainment only became significant in the engaged condition. Additionally, we found no indication of an asymmetry in the strength of delta entrainment in a previous study (Chapter 2), where monkeys were presented a rhythmic stream of pure tones and were performing a frequency deviant detection task. Thus it appears that in macaques, functional asymmetry becomes apparent when spectrotemporally more complex stimuli are used and the subjects are engaged in a task where these stimuli are relevant. In fact, human studies that show indications for functional asymmetry using electrophysiological recordings and/or neuroimaging utilized similar spectrotemporally complex, rhythmic stimuli (Boemio et al., 2005; Jamison et al., 2006; Giraud et al., 2007; Obleser et al., 2008; Morillon et al., 2010), even though in some of these studies stimuli were presented in a passive condition. We speculate that in humans, hemispheric asymmetry might be structurally more solidified via evolution, which is why functional differences can be revealed even in a passive state. Another difference between human findings and our results in monkeys is that while in our data, entrainment on both long and short time-scales was left lateralized, most human studies find that slower (delta-theta) modulations of neuronal activity related to the temporal structure of the acoustic input are lateralized to the right hemisphere. We speculate that one reason for the difference might be stronger evoked type responses at the rate of stimulation in the right hemisphere, which would bias both neuroimaging and electrophysiological measurements. Although a thorough analysis of this proposition is beyond the scope of the present study, the spectrograms in Figure 7C do provide some support for this notion: while spontaneous low frequency oscillatory amplitude is smaller in right hemisphere recordings, the amplitude increase related to auditory stimuli in the delta band is larger. It will be important to conduct human experiments with near threshold auditory stimuli in which case the effect of evoked type responses should be negligible.
Nevertheless, despite some discrepancies, our results provide strong support that similar to humans, relevant spectrotemporally complex rhythmic stimuli are processed asymmetrically by the left and right hemispheres, indicating that the functional-anatomical precursor to the machinery that enables speech perception and production is present in non-human primates, at least at lower cortical stages.

4.4.5. Conclusions

Our findings indicate that attended broadband stimuli organized on multiple timescales (i.e. repetitive click trains) result in a multi-scale entrainment of ongoing oscillations across all of A1, and that the phases of entrainment of low and high frequency oscillations are independent of each other. Nonetheless, the intricate combination of low and high excitability phases in differently tuned neuronal ensembles results in a predominantly suppressive effect on auditory responses to click trains, except in a subset of high frequency regions of A1. We hypothesize that the opposite sign excitability modulation of high vs. low frequency representation related to acoustic edges could set the stage for the predictive processing of alternatingly occurring high vs. low frequency elements of complex acoustic stimuli like speech. Additionally, evidence of superior phase consistency of entrained oscillations in left A1 provides support for functional hemispheric asymmetry even at the earliest auditory cortical processing stage and remarkably even in non-human primates.

4.5. EXPERIMENTAL PROCEDURES

4.5.1. Subjects. In the present study, we analysed the electrophysiological data recorded during 48 penetrations of area A1 of the auditory cortex of 2 female macaques (Macaca mulatta)
weighing 4-7 kg, who had been prepared surgically for chronic awake electrophysiological recordings. Prior to surgery, each animal was adapted to a custom fitted primate chair and to the recording chamber. All procedures were approved in advance by the Animal Care and Use Committee of the Nathan Kline Institute.

4.5.2. Surgery. Preparation of subjects for chronic awake intracortical recording was performed using aseptic techniques, under general anesthesia, as described previously (Schroeder et al., 1998). The tissue overlying the calvarium was resected and appropriate portions of the cranium were removed. The neocortex and overlying dura were left intact. To provide access to the brain and to promote an orderly pattern of sampling across the surface of the auditory areas, cilux recording chambers (Crist Instruments) were positioned normal to the cortical surface of the superior temporal plane for orthogonal penetration of area A1, as determined by a pre-implant MRI. Together with socketed Plexiglas bars (to permit painless head restraint), they were secured to the skull with orthopedic screws and embedded in bone cement. A recovery time of minimum six weeks was allowed before the animal was head restrained and we began data collection.

4.5.3. Electrophysiology. Animals sat in a primate chair in a dark, isolated, electrically shielded, sound-attenuated chamber with head fixed in position, and were monitored with infrared cameras. Laminar profiles of neuroelectric activity were obtained simultaneously from left and right hemisphere auditory cortices using linear array multi-contact electrodes (23 contacts, 100 µm intercontact spacing). Multielectrodes were inserted acutely through guide tube grid inserts, lowered through the dura into the brain, and positioned such that the electrode channels would span all layers of the cortex (Fig. 2), which was determined by inspecting the laminar response profile to binaural broadband noise bursts. Neuroelectric signals were impedance matched with a pre-amplifier (10x gain, bandpass dc-10 kHz) situated on the
electrode, and after further amplification (500x) they were recorded continuously in a 0.01 - 8000 Hz passband digitized at a sampling rate of 20 kHz and precision of 16-bits using custom made software in Labview. The signal was split into the field potential (0.1-300Hz) and MUA (300-5000Hz) range by zero phase shift digital filtering. MUA data was also rectified in order to improve the estimation of firing of the local neuronal ensemble (Legatt et al., 1980). One-dimensional current source density (CSD) profiles were calculated from the local field potential profiles using a three-point formula for the calculation of the second spatial derivative of voltage (Freeman and Nicholson, 1975). The advantage of CSD analysis is that CSD signals are not affected by volume conduction like the local field potentials, and they also provide a more direct index of the location, direction, and density of the net transmembrane current flow (Mitzdorf, 1985; Schroeder et al., 1998). At the beginning of each experimental session, after refining the electrode position in the neocortex, we established the best frequency (BF) of the recording site using a “suprathreshold” method (Steinschneider et al., 1995; Lakatos et al., 2005a). The method entails presentation of a stimulus train consisting of 100 random order occurrences of a broadband noise burst and pure tone stimuli presented at 50 dB loudness with frequencies ranging from 353.5 Hz to 32 kHz in half octave steps (duration: 100 ms, r/f time: 5 ms, SOA = 624.5). Auditory stimuli for tonotopy and for the behavioural task were generated at 100 kHz sampling rate in Labview using a multifunction data acquisition device (National Instruments DAQ USB-6259), and presented through SA1 stereo amplifiers coupled to FF1 free field speakers. Loudness was calibrated using measurements made with an ACO Pacific PS9200/4012 calibrated microphone system.

4.5.4. Behavioral task and stimuli. The goal of the present set of experiments was to examine the effect of engagement in an auditory task on the entrainment of neuronal oscillations on multiple time-scales, and on the auditory responses. We presented the subjects rhythmic
streams of click trains (Fig. 1A): the click trains consisted of 5 clicks (40 or 50 dB loudness), generated by driving the speakers with 0.1 ms square waves that were arranged 30.3 ms apart. The click trains were presented with a 624.5 ms constant SOA. In this rhythmic stream of standard, frequently presented click trains deviant click trains occurred at 2 – 6 s random time intervals. Deviant click trains only differed in their temporal structure: the 3rd click was delayed by 15 – 30.3 ms depending on the subject’s performance which we tried to keep between 60 – 80 % correct. To engage the monkeys in detecting deviant or target click trains, in the beginning of training, 0.25-1 ml juice reward was delivered to them simultaneously with each deviant through a tube. The tube was positioned such that the monkeys had to stick out their tongue in order to get the juice. Licking was monitored using a simple contact detector circuit (Slotnick, 2009), the output of which was continuously recorded together with the timing of standard and deviant tones for offline analyses via a multifunction data acquisition device (National Instruments DAQ USB-6259) in Labview. In this phase of training the third click in the deviant click trains was shifted by 30.3 ms corresponding to a missing 3rd click. After 2 sessions, the juice reward was omitted on every 10th deviant. If the monkeys licked on these deviants without a paired juice reward, signalling that they were engaged in the auditory task, we omitted the reward on 20% of the deviants, and also gradually decreased the shift of the 3rd click when the monkey’s performance increased to around 80%. In one of the subject’s the shift was decreased to 15 ms in the last experiments, while in the other monkey it was never below 25 ms. We only analysed data related to standard stimuli that preceded deviants (with or without juice) on which the subjects licked. Further, we only analysed CSD and MUA data related to standards that followed the deviant by a minimum of 4 stimulus positions, to avoid artifacts related to licking and to ensure that subjects re-engaged in the task (deviants could not occur for 2 seconds following a deviant/target).
Besides the engaged, auditory task condition, we recorded data during the presentation of the same stimuli in a passive condition, when the juicer was removed, and the subjects had no auditory or other task, but were quietly sitting in the recording chamber. Following the passive condition, we also recorded 3-5 minutes of spontaneous neuroelectric activity in the absence of stimuli presented.

4.5.5. Data analysis. Data were analysed offline using native and custom-written functions in Matlab (Mathworks, Natick, MA). After selective averaging of the CSD and MUA responses to the tones presented in the suprathreshold tonotopy paradigm, recording sites were functionally defined as belonging to AI or belt auditory cortices based on the sharpness of frequency tuning, the inspection of the tonotopic progression across adjacent sites, and relative sensitivity to pure tones versus broad-band noise of equivalent intensity (Merzenich and Brugge, 1973; Rauschecker et al., 1997; Lakatos et al., 2005a). In the present study only recordings obtained from area A1 were analysed. At the end of each animal’s experimental participation, functional assignment of the recording sites was confirmed histologically (Schroeder et al., 2001).

Utilizing the BF-tone related laminar CSD profile, the functional identification of the supragranular, granular and infragranular cortical layers in area A1 (Fig. 3) is straightforward based on our earlier studies (Schroeder et al., 1998; Schroeder et al., 2001; Lakatos et al., 2005a; Lakatos et al., 2007). In the present study we focused the analyses of ongoing and event related neuronal activity on the supragranular CSD with largest BF tone related activation (sink), and the MUA averaged across all layers. The reason for this selection is that both ongoing and entrained oscillatory activity are most prominent in the supragranular layer (Lakatos et al., 2005b; Lakatos et al., 2007; Lakatos et al., 2008), and they appear to reflect synchronous excitability fluctuations of the local neuronal ensembles across all layers, as
evidenced by synchronous MUA amplitude fluctuation across the layers (O'Connell et al., 2011). Also, dominant delta frequency neuronal activity in all cortical layers is largely coherent with supragranular delta oscillatory activity, with varying but stable phase differences across cortical depths (Lakatos et al., 2005b; O'Connell et al., 2011).

To determine MUA response onset latencies, the MUA averaged across all cortical layers was used, and response onset was defined as the earliest significant (> 2 standard deviation units) deviation of the averaged waveforms from their baseline (-50 – 0 ms), that was maintained for at least 5 ms.

For the analysis of ongoing and event related (entrained) delta and gamma oscillatory activity, instantaneous power and phase in single trials were extracted by wavelet decomposition (Morlet wavelet) on 345 scales from 0.5 to 55 Hz. Oscillatory amplitudes were measured in spontaneous recordings, and also in data recorded during stimulus presentation. In both cases, a continuous wavelet transform on the entire recording was performed, but in the latter case only time-points during and following the presentation of standard tones (see above) were averaged. To characterize delta and gamma phase distributions related to stimulus presentations (trials), the wavelet transformed data was normalized (unit vectors), the data at corresponding time-points relative to each stimulus onset were averaged, and the length (modulus) of the resulting vector was computed (e.g. Lakatos et al., 2007). The value of the mean resultant length, also called inter-trial coherence (ITC) ranges from 0 to 1; higher values indicate that the observations (oscillatory phase at a given time-point across trials) are clustered more closely around the mean (i.e. phase distribution is biased) than lower values. Phase distributions were evaluated statistically using circular statistical methods. Significant deviation from uniform (random) phase distribution was tested with Rayleigh’s uniformity test. Pooled
phase distributions were compared by a nonparametric test for the equality of circular means (Fisher, 1993; Rizzuto et al., 2006). The $\alpha$ value was set at 0.01 for all statistical tests.

Independent of their waveform shape (frequency composition in the frequency domain), cyclically occurring events like the suprathreshold, “evoked type” response waveforms can artificially bias phase measures at the frequency that corresponds to the stimulus presentation rate (Lakatos et al., 2013a; Zoefel and Heil, 2013). Since in some cases, visual inspection revealed a clear “evoked type” transient waveform in the supragranular CSD in response to the click train (Fig. 3), similar to our earlier studies in the case of responses to pure tones, we applied a linear interpolation to the single trials in the time interval of the evoked-type auditory response (5 – 150 ms), and determined delta phases in the interpolated data at click train onset. For the same reason, we determined gamma phases at the time of the $4^{th}$ click (90.9 ms) rather than at stimulus onset, plus gamma entrainment most likely only develops after the 3rd click.
GENERAL DISCUSSION

The three studies comprising this dissertation investigated the role of ongoing (or spontaneous) neuronal oscillations (electrophysiological context) in the modulation of stimulus related evoked responses (sensory content) in A1 during attentive and non-attentive conditions.

The findings of the first study (Chapter 1) verified our predictions: we established that one of the mechanisms of inhibitory responses (signaled by a poststimulus MUA suppression) to non-preferred frequency (non-BF) tones (~ 2 octaves different from the preferred frequency (BF) tone) is the resetting of dominant ongoing oscillations to their low excitability phases. This modulatory inhibitory response appears to compliment the effects of the previously described feedforward inhibition (Wehr and Zador, 2003; Zhang et al., 2003; Tan et al., 2004; Wehr and Zador, 2005; Wu et al., 2008), which is due to specific thalamocortical (TC) inputs targeting Layer 4 inhibitory neurons. In contrast to this, an evoked excitatory response, again the result of specific TC inputs this time synapsing on glutamatergic neurons (Miller et al., 2001; Hackett et al., 2007; Liu et al., 2007), BF tone related inputs reset supragranular ongoing oscillations to opposing, high excitability phases. We concluded that this mechanism of frequency specific phase reset in passively behaving animals could be a dynamic one that is responsible for the top down modulation of auditory information, which is transmitted by the more obligatory or “hardwired” lemniscal system (specific TC projections resulting in evoked type responses). This notion is supported by a previous study showing that phase reset is under strong attentional control (Lakatos et al., 2009). Thus we were prompted to investigate what role this mechanism plays in enhancing the sensory representation of attended auditory stimulus streams (Chapter 2’s study).
In general the results of the second study (Chapter 2) also agreed with our predictions; we found that during an auditory attention task with rhythmic stimulation, low frequency (delta band) oscillatory activity was modulated so that it entrained to the timing of stimuli in the attended stimulus stream (i.e. entrained with a period that matched the stimulus presentation rate), and that the phase of entrainment was frequency specific: if the frequency of pure tones in the attended stream matched the BF of a given A1 site, oscillations were entrained to their high, while if the attended tone was a non-BF one, oscillations were entrained with their low excitability phases (as expected from Chapter 1's findings). Consequently, responses to attended BF tones were amplified, while responses to attended non-BF tones were suppressed compared to those in the ignored condition, leading to a sharpening of frequency tuning. Another finding of the study was that ignored auditory stimuli did not modulate ongoing oscillations, providing further evidence that phase reset and entrainment are strongly modulated by attention. Importantly, we found that attention-related entrainment to the low excitability phase was not confined to the attended tones whose BF regions neighbored the recording site (Chapter 1 and Lakatos et al., 2013a). Rather the effect extended several octaves removed from the BF. Thus we concluded that, as a recent human study predicted (Lakatos et al., 2013b), if rhythmic pure tones are attended, almost all of A1 – with the exception of the BF region – entrains its neuronal activity to their occurrence to predicatively suppress frequency content that would interfere with the attended one. This mechanism suggests a multidimensional filter role for ongoing oscillatory activity in A1 – organized across cortical space and time by phase reset and entrainment – that enhances the sensory representation of attended auditory stimulus streams along both temporal and spectral feature dimensions. Since in this study the monkeys were performing a frequency discrimination task, we additionally tested if there was evidence of hemispheric specialization for processing spectral features
(Zatorre and Belin, 2001; Zatorre et al., 2002). We found that neither the sharpening of tuning, nor delta entrainment showed a hemispheric difference.

Our conclusion from study two (Chapter 2), that ongoing neuronal oscillations can be harnessed by the brain to act as spectrotemporal filters during a rhythmic frequency discrimination task, encouraged us to examine how this mechanism would operate if spectral features (and thus the need for spectral filtering) were eliminated from the task, and only the temporal structure of stimuli remained important. Also, since most human studies that showed a hemispheric lateralization of auditory function employed spectrotemporally complex sounds, we utilized click trains organized on multiple time-scales in our third study (Chapter 3). Specifically, monkeys attended to auditory stimuli which had a broad frequency spectrum, i.e. click trains, and were arranged so that the deviants differed only in their temporal structure. Nevertheless, our prediction that attended broadband auditory stimuli would entrain oscillations with their high excitability phases across all of A1, thereby enhancing responses across the board was not supported. Rather, we discovered, to our surprise, that the main effect of attention was response suppression in most A1 regions. This was due to the entrainment of low (delta - reflecting presentation rate of the click trains) and high (gamma - reflecting the presentation rate of the clicks within the click trains) frequency oscillations with a combination of phases and amplitudes – due to phase amplitude coupling of delta – gamma oscillatory activity (Lakatos et al., 2005b) – that ultimately resulted in response suppression. However recording sites in the 11-16 kHz region of A1 either failed to exhibit this attentional response suppression or showed a response enhancement. This was linked to both delta and gamma oscillations entraining with their high excitability phases. This finding brings up the possibility that there could be functional specialization across differently tuned neuronal ensembles in A1: the function of neuronal ensembles tuned to high frequencies could be more important in the
temporal than in the spectral domain. In other words, rather than analyzing the spectral content of auditory stimuli (most behaviorally important spectral content in human and monkey vocalizations is below 8 kHz), they would transmit rapid and accurate information about the timing of acoustic landmarks. Aside from the surprising finding of mostly “suppressive entrainment”, our expectations regarding the specialization of left A1 for temporal processing were realized, as we found greater temporal precision of neuronal oscillatory alignment (entrainment) in left A1 as signified by enhanced phase locking (higher ITC values) to the stimuli in both the delta and gamma frequency bands. This finding is important since it shows that at least on the level of primary auditory cortex, macaques are an acceptable model for humans in terms of how spectrotemporally complex auditory stimuli are processed by the brain.

5.1. RELATIONSHIP OF STUDIES

Apart from the overall theme of this dissertation, which proposes a model of how ongoing neural oscillations are harnessed by the brain to reactively (through phase reset) and predictively (through entrainment) modulate neuronal excitability during auditory stimulus processing under varying stimulus and task conditions, several findings from the three chapters are interrelated. For instance, the conclusion that stimulus frequency-dependent entrainment of delta oscillatory phase modulates gamma band activity amplitudes, and thus pre-stimulus neuronal excitability across A1 (Chapter 2), and our reasoning that the delta phase dependent post-stimulus amplitude of gamma oscillations has an important effect on stimulus processing in the case of hierarchically organized rhythmic stimuli (Chapter 3), along with results of earlier studies (Lakatos et al., 2005b;Canolty et al., 2006;Lakatos et al., 2008;Canolty and Knight, 2010) highlights the significance of phase-amplitude coupling on cortical processing.
These results also emphasize the dynamic flexibility of neuronal oscillations in forming the internal neurophysiological context within which the stimulus related content is processed.

As attention to the two distinct types of stimuli in Chapter 2 and 3’s experiments (i.e. pure tones and click trains) similarly resulted in counterphase entrainment of delta oscillations across differently tuned frequency representations in A1, we hypothesize that counterphase entrainment might be the “default operational mode” in A1. This operational mode could have developed to support speech processing (or vice versa), since in speech there is an apparent counterphase structure of high vs. low frequency elements. The counterphase entrainment in the case of both stimulus type also might mean that this default mode operation is “hardwired”. If so, it may be that the predictive enhancement in excitability at one frequency representation (i.e. BF or at the altered frequency representation of a click: 11-16 kHz) through entrainment results in a predictive suppression of neuronal activity in all other A1 regions through some kind of a “winner take all” mechanism. The anatomical routes of these effects are unknown, but there are two main theoretical possibilities: one could be the horizontal modulation of excitability within supragranular A1 neuronal ensembles, while the other could be the modulation of non-specific thalamocortical inputs either via intrathalamic connections or corticothalamic feedback.

Two of our studies also provide some evidence that high frequency regions of A1 may be functionally distinct from other frequency regions of A1 (Chapter 1 & 3). In Chapter 1 we found that inhibition to non-BF tones in high frequency regions was not complete but that these sites were likely to respond with a short excitation to lower frequency tones followed by the inhibitory response. Intriguingly the BF of these sites appears to match that of the A1 regions (BF = 11-16 kHz) which fail to show response suppression to click trains in the engaged condition in Chapter 3. Therefore, it seems that in addition to preferentially processing
transient broadband stimuli (or acoustic edges) and consequently either directly or indirectly triggering a spatiotemporal excitability pattern across low and high frequency regions of A1, these sites tuned to 11-16 kHz are "alerted" to the occurrence of other frequency tones as signified by the brief excitatory MUA component before the sustained inhibition. This is in agreement with our speculation in Chapter 3 and above that high frequency A1 regions may play an important role in the perception of temporally complex stimuli, like monkey vocalizations or speech, by always transmitting information about the timing of the distinct acoustic elements or landmarks. Thus, possibly the output of these neuronal ensembles is key in the parsing and grouping of auditory information.

5.2. HEMISPHERIC LATERALIZATION OF AUDITORY FUNCTION

There are two main theories on lateralized auditory cortex specialization. The first theory posits that the auditory cortices process incoming stimuli on distinct timescales in each hemisphere as reflected by the high frequency (gamma) dominance of auditory stimulus related activity in the left, and low frequency (delta-theta) dominance in the right (Poeppel, 2003; Boemio et al., 2005; Luo and Poeppel, 2012). The second one proposes that the left auditory cortex is endowed with enhanced temporal processing abilities while the right exhibits a relatively finer resolution in the frequency domain (Zatorre and Belin, 2001; Zatorre et al., 2002). The results of the studies conducted in Chapter's 2 and 3 appear to partially support Zatorre’s rather than Poeppel’s hypothesis as we show that during a temporal task, which used broadband stimuli, left A1 displays greater stimulus structure-related oscillatory activity or in other words more precise entrainment to the stimuli in both high (gamma) and low (delta) frequency bands. This effect, while evident in the passive was significantly enhanced in the engaged condition (Chapter 3). However, during the frequency discrimination task, which used pure tones, we
found no difference in strength of entrainment of delta oscillations between the two hemispheres or in the sharpening of tuning (Chapter 2), indicating that spectrally related modulation of oscillatory activity is not enhanced in the right auditory cortex of monkeys as would be expected from Zatorre’s model. Our finding that in non-human primates functional asymmetry in A1 only becomes apparent when spectrotemporally more complex stimuli are used and the task is a temporal one, is consistent with the findings of a recent study in macaques which showed that complex auditory stimuli such as human speech preferentially activated left auditory belt areas (Joly et al., 2012). The reason that the authors may not have found functional lateralization in A1 could be due to the fact that the stimuli were not made task relevant.

5.3. NEURONAL OSCILLATIONS AND BRAIN DISORDERS

Since a major topic of this dissertation is the importance of ongoing neuronal oscillations for efficient sensory processing, it follows that impaired oscillations or oscillatory mechanisms should lead to inefficient or altered perceptual processes. Thus in this section we review some evidence that oscillatory mechanisms are involved in brain disorders, and consider how our findings can lead to a better understanding of the neurophysiological deficits that underlie brain disorders.

Above and in Chapter 3 we discussed the importance of neuronal oscillations and oscillatory entrainment in speech processing. Linked to speech processing is the learning disability developmental dyslexia, which is characterized by the inability to decode words into their phonemic segments (Lyon, 1995). Interestingly, children diagnosed with specific language impairment (i.e. weak oral language abilities but normal non-verbal capabilities) display a deficit in processing quickly presented acoustic stimuli at a gamma frequency rate,
which incidentally covers the phonemic level of speech perception (Tallal et al., 1993). Therefore it is quite possible that – as previously proposed (Goswami, 2011; Lehongre et al., 2011) – inadequate temporal sampling mediated by altered neuronal oscillations could be responsible for these disabilities. Indeed, results of a recent study show that dyslexic subjects exhibit reduced gamma entrainment to 30 Hz amplitude modulated white noise in left auditory cortex compared to normal readers (Lehongre et al., 2011), suggesting that certainly an inability of the left hemisphere to temporally track and parse speech at the correct phonemic rate could be behind this and other language impairments. However it is unclear in this population whether the generation of endogenous gamma band oscillations are abnormal (e.g. dysfunctional neurotransmitters) or whether the mechanism of phase reset could be disrupted and thus the entrainment of these oscillations.

Schizophrenia is a serious mental disorder that affects up to 1% of the population worldwide. Many studies have theorized that several of the higher order cognitive deficits that schizophrenic patients exhibit, for example inability to sustain attention (O’Grada et al., 2009), difficulty with working memory (Park et al., 1999) and inability to interpret prosody (Leitman et al., 2005) may be in fact related to deficits in early sensory processing (Javitt et al., 2000; Butler et al., 2001). This is supported by the generation of abnormal ERP components, which index low level sensory processing, such as reduced amplitude MMN and auditory N100 in schizophrenic subjects compared to healthy controls (Butler et al., 2001; Umbricht et al., 2003; O’Donnell et al., 2004; Umbricht and Krljes, 2005; Rosburg et al., 2008; Turetsky et al., 2009). Since ERP components are at least partly generated by phase resetting of ongoing oscillations (Makeig et al., 2004; Fuentemilla et al., 2006; Min et al., 2007; Barry, 2009; Telenczuk et al., 2010), it is possible to think that deficits in ERP generation (e.g. MMN and N100) observed in schizophrenics could be related to a dysfunction in phase resetting of
neuronal oscillations. This is exactly what Jansen and colleagues have shown: schizophrenic patients have a phase locking deficiency as compared to controls, particularly in the delta-theta frequency range, during the time period 40-160ms after an auditory stimulus presentation (Jansen et al., 2004). The authors found a significant correlation between N100 amplitude and degree of phase locking 100 ms post stimulus indicating that phase locking (possibly indicating phase reset) in the theta band contributes to the N100 component. Another group (Brockhaus-Dumke et al., 2008) also found reduced phase locking in the theta band as well as in the alpha (8-12Hz) band in schizophrenia patients. Similar to Jansen’s group, this was significantly correlated with the amplitude decrease of the N100 component after presentation of an auditory click.

The most direct evidence for impaired phase reset and thus entrainment comes from a recent study by Lakatos and colleagues (Lakatos et al., 2013b). This study showed that delta oscillations in healthy control subjects are entrained by attended stimuli in a task difficulty dependent manner in an auditory frequency discrimination task. However schizophrenia patients performing the same task exhibit a lack of entrainment, even though they perform the task with presumably the same effort as indicated by their performance and task related parieto-occipital alpha amplitude changes. This finding indicates that schizophrenia is characterized by the inability to align the internal, electrophysiological context to the timing of relevant external events. In terms of auditory perception of pure tone frequency, this should result in less frequency selectivity based on our results (Chapter 2), which is exactly what the Lakatos study found. The authors theorized that a lack of counter-phase entrainment across A1 is responsible for the patient’s increased tone discrimination thresholds during tone matching or auditory discrimination tasks. Besides this, based on what we know about the function of entrainment and phase reset, the inability to synchronize excitability fluctuations to the timing of inputs
should result in instable and distorted perception, which is one of the hallmark symptoms of schizophrenia.

5.4. FUTURE DIRECTIONS

Since non-primary auditory regions have degraded tonotopic representations (Kaas and Hackett, 2000), it is quite likely that the oscillatory mechanism which acts as a spectrotemporal filter during auditory selective attention tasks in A1 (Chapter's 2 and 3), plays only little if any role in the attentional selection of relevant items in belt areas. We propose that the main mechanisms that further enhances the sensory representation of relevant, while suppressing the representation of irrelevant items in belt is the alignment of ongoing oscillatory activity in belt to oscillatory activity in A1 regions that process attended stimulus features. This would only allow efficient downstream communication at times when relevant events occur (communication through coherence (Fries, 2005), while higher order auditory processing regions (belt and parabelt cortical areas) would be relatively “closed” to communication at all other times. So for example, core and belt regions processing the same frequency would be in a similar phase (e.g. high excitability phase) on a time scale that corresponds to the temporal structure of attended stimuli. This proposition can be tested in experiments utilizing simultaneous A1-belt recordings. Another, related question is whether phase reset that aligns oscillations for efficient communication (functional connectivity) occurs through hierarchically coupled regions sequentially, or if a common input resets the phase of ongoing oscillatory activity in all auditory cortical regions. This again could be tested in data recorded with paired core-belt recordings while the monkey performs a similar intermodal attention task as we used in Chapter 2: if there is some delay in the timing of maximal phase locking (ITC) of ongoing
oscillatory activity following an auditory or visual stimulus, this would indicate sequential activation, while no delay would indicate simultaneous phase reset by common inputs.

As we speculated above and in Chapter 1 and 2's discussions, the mechanism of frequency specific phase reset and thus entrainment of neuronal oscillations is probably due to activation of the supragranular layers by nonspecific (nonlemniscal) thalamocortical afferents that most likely originate in the medial region of the medial geniculate nucleus (MGNm). This is supported by findings that show that pure phase reset responses are initiated in the supragranular layers and are heavily weighted towards the supra- and somewhat towards the infragranular layers, skipping the granular layer (Chapter 1; Lakatos et al., 2007), which matches the anatomical connectivity pattern of the nonspecific thalamic system (Molinari et al., 1995; Jones, 2001). Also, the activation (phase reset) of primary auditory and visual cortical areas by attended auditory or visual stimuli occurs at the same time (Lakatos et al., 2009), making a lateral or feedback type of influence of one modality by the other unlikely. To confirm our theory of the involvement of non-specific thalamus in phase reset and entrainment, it would be useful to conduct paired MGNm-A1 recordings while the monkey performs a rhythmic frequency discrimination task, and inactivate the MGNm through direct application of lidocaine or muscimol (a GABA agonist). We expect that supragranular oscillations which had become entrained to the presentation rate of the attended tones would cease entrainment after pharmacological inactivation of non-specific auditory thalamus regions. Also, the behavioral effect of MGNm inactivation might be seen as a reduction in target tone detection rate and increased tone discrimination thresholds similar to what is seen in schizophrenic patients (Lakatos et al., 2013b). On the flip side, we envision that rhythmic electric microstimulation of the same thalamic regions that is synchronized to the timing of near threshold auditory stimuli would increase task performance.
In conclusion, we envision that deciphering the mechanisms that govern phase reset and entrainment will further our mechanistic understanding of the basis of developmental and neuropsychiatric disorders that exhibit a deficiency or corruption of these essential modulatory brain mechanisms.
REFERENCES


