Magnetoencephalography and Auditory Neural Representations

J.Z. Simon[1,2] and N. Ding[1]

[1] Department of Electrical & Computer Engineering, University of Maryland, College Park MD 20815, USA
[2] Department of Biology, University of Maryland, College Park MD 20815, USA

Abstract— **Complex sounds, especially natural sounds, can be parametrically characterized by many acoustic and perceptual features, one among which is temporal modulation. Temporal modulations describe changes of a sound in amplitude (amplitude modulation, AM) or in frequency (frequency modulation, FM). AM and FM are fundamental components of communication sounds, such as human speech and species-specific vocalizations, as well as music. Temporal modulations are encoded in at least two ways, temporal coding and rate coding. Magnetoencephalography (MEG), with its high temporal resolution and simultaneous access to multiple auditory cortical areas, is a non-invasive tool that can measure and describe the temporal coding of auditory modulations. We refer to the neural temporal encoding of temporal acoustic modulations as "modulation encoding". For simple, individually presented, acoustic modulations, modulation encoding is well described by a simple modulation transfer function (MTF). Even in this simple case, however, the MTF may depend strongly on the type of modulation being encoded (e.g. AM vs. FM, narrowband vs. broadband) or the context in which the modulation is heard (e.g. attended vs. unattended). Here we present a range of different types of modulation encoding employed by human auditory cortex. The simplest examples are for sinusoidally amplitude modulated carriers of a range of bandwidths (with special emphasis on those modulation rates relevant to speech and other natural sounds: below a few tens of Hz). We provide evidence that the modulation transfer functions are lowpass in shape and relatively independent of bandwidth. When several modulations are applied concurrently however, the modulation encoding typically, but not always, becomes non-linear: the auditory modulations are at the rates of the acoustic modulations but also at the rates of cross-modulation frequencies. The physiological occurrence, or not, of these cross terms seem be in accord with the psychophysical concept of modulation filterbanks.**

Keywords— **Magnetoencephalography, cortex, modulations.**

I. Introduction

Speech and environmental sounds contain slow modulations in both amplitude (AM) and frequency (FM). Speech intelligibility depends strongly on the integrity of modulations within a range of perceptually-relevant rates, the same range which drives cortical responses most vigorously as measured by single-unit activity. The neural mechanisms by which this modulation is encoded are important to our understanding of perception, and to applications such as auditory prostheses. Speech and natural sounds contain simultaneous AM and FM, of various rates and embedded in other sounds, so it is crucial to determine the encoding of compound and noisy modulations as well as simple ones.

In humans, neurophysiological studies of temporal methods in auditory coding are generally limited to non-invasive techniques. For these studies, MEG is an appropriate tool, as it has excellent temporal resolution, is silent, and can localize neural sources with suitable resolution (particularly so in auditory areas).

A. Modulation Sensitivity

Psychophysically, human listeners attending to spectrally simple broadband sounds are most sensitive to amplitude modulations at a few Hz, peaking near 10 Hz and diminishing substantially by 100 Hz [1]. This result also holds when increasing the spectral complexity of the carrier (see, e.g. [2,3]). Speech intelligibility is strongly dependent on modulation rates below 20 Hz, for both amplitude modulation [4,5] and frequency modulation components of speech [6].

B. MEG and Temporal Response Properties

Complementary to hemodynamic methods (e.g. fMRI), MEG is more limited in its spatial resolution but virtually unlimited in its temporal resolution (e.g. down to ~1 ms). Since temporal aspects of auditory signals play such a crucial role in both perception and speech processing, MEG is a logical tool to employ for these purposes. Additionally, MEG is silent, unlike fMRI.

MEG measures the magnetic fields generated by neuronal current flow, using SQUID based detectors (Superconducting Quantum Interference Devices) to make high-gain recordings [7]. It is believed the main source for MEG is current flow in pyramidal cells' apical dendrites. [7]

A major benefit over EEG (which measures electric potentials generated by the same neural activity) is high sensitivity to activity within auditory cortex (where pyramidal cell dendrites are parallel to the surface). In contrast, EEG responses mix across cortical hemispheres and are strongest medially (see, e.g., [8]). Another benefit to MEG is that results from data averaged over subjects are typically visible in individual subjects as well.

K.E. Herold, W.E. Bentley, and J. Vossoughi (Eds.): SBEC 2010, IFMBE Proceedings 32, pp. 45–48, 2010.
www.springerlink.com

The auditory steady state response (SSR) is the neural response to a stationary acoustic stimulus. SSR responses are a rich source of neurophysiological data that have strong potential to make compelling links with animal-based cortical physiology. Auditory SSR latencies from MEG, typically 30-40 ms [9,10], are much better matched to latencies seen in single neuron studies than the M100 or even M50.

SSR was first measured in humans by using EEG, where the strongest response is typically found at 40 Hz [11]. Rates below 20 Hz are less well explored due the lower signal to noise ratio (SNR) compared to near 40 Hz. Using MEG, Ross et al. [9] found a slowly decreasing plateau from 20 Hz down to 10 Hz, but with increasing error bars. EEG studies give conflicting results, but, with care in rejecting strong but statistically non-significant responses, Picton et al. [8] found a trend that the EEG response increases as rate decreases below 20 Hz. Our MEG results support this result in much greater detail.

II. Methods

15 subjects (9 female) (Experiment 1), and 10 subjects (5 female) (Experiment 2), all right handed [12], who reported normal hearing and no history of neurological disorder, listened to the acoustic stimuli while MEG recordings were taken. The procedures were approved by the University of Maryland institutional review board and written informed consent was obtained from each participant. Subjects were paid for their participation.

In Experiment 1, sinusoidally amplitude-modulated sounds of 2000 ms duration were presented to each subject. Twenty stimuli were made with 5 modulation frequencies (1.5 Hz, 3.5 Hz, 7.5 Hz 15.5 Hz and 31.5 Hz) and 4 different carriers (pure tone at 707 Hz; 1/3 octave pink noise; 1 octave pink noise Hz and 5 octave pink noise (all centered at 707 Hz) with 100% modulation depth. The four different bandwidth conditions are henceforth referred to as 0, 1/3, 1, and 5 octave bandwidths.

In Experiment 2, the stimulus is a tone simultaneously frequency and amplitude modulated. The FM modulation frequency, f_{FM}, is fixed at 37.7 Hz. The maximum frequency deviation from the base carrier frequency, 550 Hz, is 330 Hz. The AM modulation frequency, f_{AM}, is varied, taking on values of 0.3, 0.7, 1.7, 3.1, 4.9, 9.0, and 13.8 Hz. The specific values of f_{FM} and f_{AM} are selected to avoid harmonic overlaps. Stimulus duration is 21 s.

Subjects were placed horizontally in a dimly lit magnetically shielded room (Yokogawa Electric Corporation, Tokyo). The signals were delivered to the subjects' ears with 50 Ω sound tubing (E-A-RTONE 3A, Etymotic Research, Inc), attached to E-A-RLINK foam plugs inserted into the ear-canal and presented binaurally at a comfortable loudness of approximately 70 dB SPL.

MEG recordings were conducted using a 157-channel axial gradiometer whole-head system (Kanazawa Institute of Technology, Kanazawa, Japan). Its detection coils form a uniform array on a helmet-shaped surface of the dewar bottom, with about 25 mm between the centers of two adjacent 15.5 mm diameter coils. Sensors are axial gradiometers with 50 mm baseline; field sensitivities are 5 fT/√Hz or better in the white noise region. Three magnetometers measure the environmental magnetic field. The signals were bandpassed between 0 or 1 Hz and 200 Hz, notch filtered at 60 Hz, and sampled at the rate of 500 Hz or 1kHz. The 1 Hz high pass filter's influence on the amplitude and phase of MEG recordings is corrected.

Two denoising techniques were applied off-line: TS-PCA [13], to remove external noise (filtered versions of the reference signals), and SNS [14], to remove noise internal to individual gradiometers. TS-PCA used a ±100 ms range of filter taps; SNS used 10 channel neighbors. Finally, DSS [15] a blind source separation technique was applied to preserve phase-locked neural activities. The DSS components are sorted based on how much of the response power is phase-locked to the stimulus. Only the first component is kept for further analysis in this study.

A single current dipole modeled the auditory response in each hemisphere. The lead field was calculated using the complex version of the Sarvas spherical head model [16]. A isotropic sphere model was build for each subject, using MEG Laboratory 2.001M (Yokogawa Electric Corporation, Tokyo). The dipole moment is estimated using least squares [17] while the dipole position is estimated using a modified simplex search with clustering [18].

III. Results

A. Experiment 1

The relation between the strength of neural AM response and the stimulus f_{AM} is commonly known as modulation transfer function (MTF). Dipole strengths of MEG responses are shown in Fig. 1 as a function of the stimulus AM rate and carrier bandwidth, for each hemisphere. This low-pass pattern is also seen in intracranial recordings [19]. The MEG response power in the right hemisphere is stronger than that in the left hemisphere for 1.5 Hz AM with pure tone carrier and 31.5 Hz AM with all carriers (paired t-test, $t(13) > 2.2$, $p < 0.05$). The MEG response power is significantly affected by the stimulus AM rate (2-way ANOVA, $F(4, 279) > 22$, $p < 10^{-4}$ for both hemispheres) but not the stimulus bandwidth. There is a significant interaction between the effect of AM rate and the effect of carrier

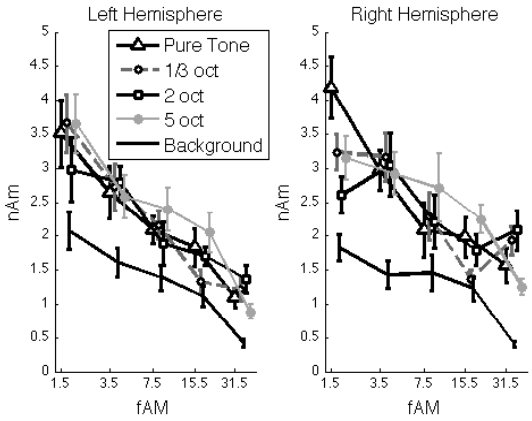

Fig. 1 Dipole strengths of MEG responses averaged over subjects. Error bars are standard error over subjects

bandwidth ($F(12, 279) > 2.04$, $p < 0.03$) in the right hemisphere.

B. *Experiment 2*

An MEG response at the stimulus f_{AM} is observed in all stimulus conditions.. The MTF measured by the power of the MEG response at f_{AM} has a low-pass pattern: the power of the MEG response to an AM sound decreases with increasing f_{AM} of that sound. It needs to be clarified, however, whether the low-pass pattern of the MTF results from stimulus-driven SSR or background noise. One can estimate the power of stimulus-driven SSR by subtracting the estimated power of background noise at f_{AM} from the power of measured MEG signal at f_{AM}. The MTF measured by this corrected power of the SSR at f_{AM} still shows a low-pass pattern and can be modeled as a linear function of f_{AM} measured in Hz (Fig. 2). The slope of the fitted linear function is -0.96 dB/Hz (99% confidence interval, -1.17 to -0.74 dB/Hz). For f_{AM} higher than 1 Hz, the slope of the MTF can also be fitted as -3.6 dB/oct (99% confidence interval, -4.8 to -2.5 dB/oct). Since the slope of the fitted line is significantly negative ($p < 0.01$), the low-pass pattern of the MTF is statistically significant for f_{AM} lower than 15 Hz. To reduce subject-to-subject variability, the corrected power is normalized before the regression analysis. Even without any correction or normalization, the slope of the MTF is still significantly negative (99% confidence interval, -1.73 dB/Hz to -0.29 dB/Hz). To investigate whether the reduction in the evoked power of the MEG response at f_{AM} is due to a loss of energy in every single trial or a loss of phase locking over trials, we calculated the phase coherence value [20] of the MEG response at f_{AM} over trials. One way ANOVA shows the phase coherence values does not significantly change when the stimulus f_{AM} increases from 0.7 Hz to 13.8 Hz ($F(5,54) = 0.84$, $p > 0.5$). Hence, the

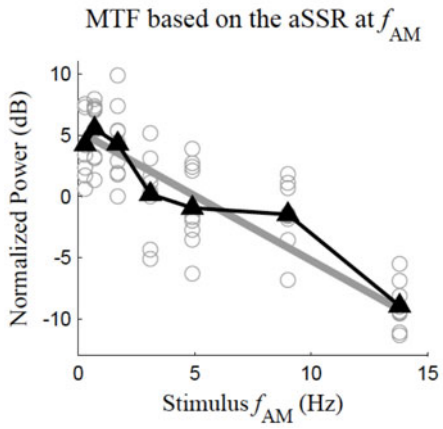

Fig. 2 Analysis of the power of the SSR at f_{AM} . The MTF calculated as a function of the corrected power of the SSR at f_{AM} and stimulus f_{AM}. Each gray hollow circle represents the corrected power of the SSR at f_{AM} for one subject. The black line marked by triangles shows the grand averaged MTF. The gray line is the optimal linear fit of the MTF

low-pass pattern of evoked power of SSR at f_{AM} is due to a change in single trial power rather than a change in over trial phase coherence. Since both the neural response power and background noise power are strongest at low frequencies, regression analysis was used to show that the signal to noise ratio of the neural response at f_{AM} does not significantly increase or decrease when f_{AM} increases ($p > 0.6$). If the neural response power at f_{AM}, $2f_{AM}$ and $3f_{AM}$ are combined, the MTF has a slope of -1.06 dB/Hz (99% confidence interval, -1.30 dB/Hz to -0.82 dB/Hz).

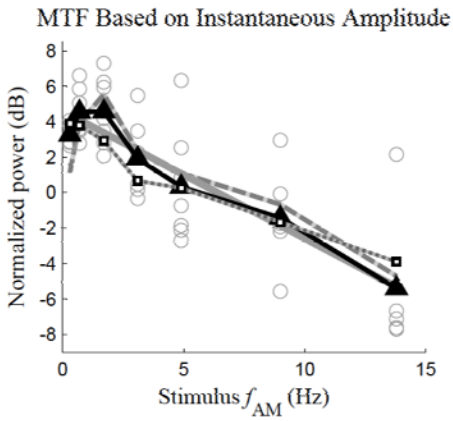

Fig. 3 Analysis of the instantaneous amplitude of the SSR at f_{FM} . The solid black line with triangle markers is the MTF averaged over all subjects. The solid gray line is the optimal linear fit of the MTF while the dotted gray line with white square markers is the MTF predicted a model. The power of the instantaneous amplitude for each subject and each condition is shown as a gray hollow circle. The instantaneous amplitude's power at f_{AM} is plotted as the dashed gray line

One of the primary goals of this work is to examine the interaction between fast modulations and slow modulations. Since the instantaneous amplitude of the SSR at f_{FM} oscillates with fundamental frequency f_{AM}, it is a neural correlate of the stimulus slow AM. Consequently, the relation between the power of the instantaneous amplitude and the stimulus f_{AM} can also be regarded as an effective MTF. We estimate the power of the instantaneous amplitude as the sum of the power at the first four harmonics of f_{AM}. This MTF (Fig. 3) has a slope of -0.72 dB/Hz (99% confidence interval, -0.95 to -0.49 dB/Hz). When f_{AM} is higher than 1 Hz, the slope of the MTF can also be fitted as -3.0 dB/oct (99% confidence interval, -5.0 to -1.6 dB/oct). For this MTF calculation, the power at each harmonic of f_{AM} was corrected by subtracting the power of background noise at that frequency; the estimate of the power of the instantaneous amplitude is also normalized to reduce subject-to-subject variability Without any correction or normalization, the MTF slope is still significantly negative ($p < 0.01$).

As the stimulus f_{AM} increases, the power of the MEG response at f_{FM} decreases at 0.86 dB/oct while the power of the instantaneous amplitude of the SSR at f_{FM} decreases 3.0 dB/oct. If the SSR at f_{FM} is assumed to be sinusoidally amplitude modulated, the neural AM modulation depth of the SSR can be estimated based on the ratio between the power of the instantaneous amplitude of the SSR and the power of the MEG response at f_{FM} . Hence, with the sinusoidal AM assumption, the neural AM modulation depth should decreases at 2.1 dB/oct.

IV. Conclusions

First, this study characterizes the properties of MEG responses to AM below 30 Hz. The SSR is strongest at the lowest modulation rates and decreases 2-4 dB per octave. For jointly modulated stimuli, the instantaneous amplitude of the SSR at f_{FM} also oscillates with fundamental frequency f_{AM}. Due to these neural interactions, the information in slow AM is simultaneously encoded in neural oscillations at f_{AM} and f_{FM} .

Acknowledgments

We thank Max Ehrman and Jeff Walker for excellent technical support. This research was supported by the National Institutes of Health (NIH) grant R01DC008342.

References

1. Viemeister NF (1979) Temporal modulation transfer functions based upon modulation thresholds. J Acoust Soc Am 66:1364-1380
2. van Zanten GA, Senten CJ (1983) Spectro-temporal modulation transfer function (STMTF) for various types of temporal modulation and a peak distance of 200 Hz. J Acoust Soc Am 74:52-62
3. Chi T, Gao Y, Guyton MC, Ru P, Shamma S (1999) Spectro-temporal modulation transfer functions and speech intelligibility. J Acoust Soc Am 106:2719-2732
4. Steeneken HJ, Houtgast T (1980) A physical method for measuring speech-transmission quality. J Acoust Soc Am 67:318-326
5. Drullman R, Festen JM, Plomp R (1994) Effect of temporal envelope smearing on speech reception. J Acoust Soc Am 95:1053-1064
6. Zeng FG, Nie K, Stickney GS, Kong YY, Vongphoe M, Bhargave A, Wei C, Cao K (2005) Speech recognition with amplitude and frequency modulations. Proc Natl Acad Sci U S A 102:2293-2298
7. Hamalainen M, Hari R, Ilmoniemi RJ, Knuutila J, Lounasmaa OV (1993) Magnetoencephalography - Theory, Instrumentation, and Applications to Noninvasive Studies of the Working Human Brain. Reviews of Modern Physics 65:413-497
8. Picton TW, John MS, Dimitrijevic A, Purcell D (2003) Human auditory steady-state responses. Int J Audiol 42:177-219
9. Ross B, Borgmann C, Draganova R, Roberts LE, Pantev C (2000) A high-precision magnetoencephalographic study of human auditory steady-state responses to amplitude-modulated tones. J Acoust Soc Am 108:679-691
10. Schoonhoven R, Boden CJ, Verbunt JP, de Munck JC (2003) A whole head MEG study of the amplitude-modulation-following response: phase coherence, group delay and dipole source analysis. Clin Neurophysiol 114:2096-2106
11. Galambos R, Makeig S, Talmachoff PJ (1981) A 40-Hz auditory potential recorded from the human scalp. Proc Natl Acad Sci U S A 78:2643-2647
12. Oldfield RC (1971) The assessment and analysis of handedness: The Edinburgh inventory. Neuropsychologia 9 (1), 97-113
13. de Cheveigné A, Simon JZ (2007) Denoising based on time-shift PCA. J. Neurosci. Methods 165 (2), 297-305
14. de Cheveigné A, Simon JZ (2008a) Sensor noise suppression. J. Neurosci. Methods 168 (1), 195-202
15. de Cheveigné A, Simon JZ (2008b) Denoising based on spatial filtering. J. Neurosci. Methods 171 (2), 331-339
16. Sarvas J (1987) Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem. Phys. Med. Biol. 32, 11-22
17. Mosher JC, Baillet S, Leahy RM (2003) Equivalence of linear approaches in bioelectromagnetic inverse solutions. IEEE Workshop on Statistical Signal Processing, St. Louis
18. Uutela K, Hamalainen M, Salmelin R (1998) Global optimization in the localization of neuromagnetic sources. IEEE Trans. Biomed. Eng. 45 (6), 716-723
19. Liegeois-Chauvel C, Lorenzi C, Trebuchon A, Regis J, Chauvel P (2004) Temporal Envelope Processing in the Human Left and Right Auditory Cortices. Cereb Cortex 14:731-740
20. Fisher NI (1993) Statistical analysis of circular data. Cambridge [England] ; New York, NY, USA: Cambridge University Press

Author: Jonathan Z. Simon
Institute: Electrical & Computer Engineering
Street: University of Maryland
City: College Park, MD 20815
Country: USA
Email: jzsimon@umd.edu