Cortical Representations of Compound Temporal Modulations: Implications for Modulation Filter Models

Juanjuan Xiang[1], David Poeppel[2,3,4], Jonathan Z Simon[1,3,4]
[1]Department of Electrical & Computer Engineering, [2]Department of Linguistics ,
[3]Department of Biology, [4]Neuroscience and Cognitive Science Program

University of Maryland College Park

Computational Sensorimotor Systems Lab

Introduction

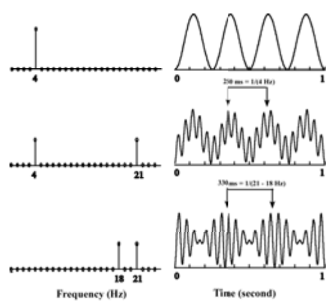
Natural sounds, including animal vocalizations and human speech, are time varying signals. Information contained in the dynamic temporal structure is crucial for speech recognition, pitch processing and stream segregation. To investigate auditory encoding of temporal information, sinusoidally amplitude modulated (SAM) stimuli and electroencephalography (EEG) and magnetoencephalography (MEG), powerful tools to explore the timing of neural processes at the level of cell assemblies, have been widely used in physiological studies. The EEG/MEG signals evoked by modulated sounds are characterized by the auditory steady state response (aSSR), a spectral component at the same frequency as the stimulus modulation frequency.

In typical situations sounds contain concurrent modulations. The ability of auditory cortex to represent multiple simultaneous modulations is critical for everyday tasks such as speech perception, where both syllabic rate (~4 Hz) and phonemic rate (~ 20 Hz) need to be tracked and processed simultaneously. To investigate the auditory encoding of temporal information in complex sounds, we employ SAM stimuli containing both single and compound modulations and measure the neural responses to these modulated stimuli by using MEG, in order to address three questions: First, how are concurrent amplitude modulations represented in the auditory cortex? Secondly, how do the neural representations of concurrent modulations fit into the current debate on modulation filters? Thirdly, what is neural correlations underlying the perception of dual modulated stimulus? The results support the existence of modulation filter banks whose filter bandwidths are band-limited (with respect to modulation rates), allowing syllabic and phonemic modulations to be processed separately. The neural underlying correlated with the perception of dual modulated stimulus is also established.

Methods

- Stimuli
  - 10 stimuli, 50.25s duration; inter-stimulus intervals from 1.5 to 1.9 s; loudness 70 dB SPL
  - 16 subjects listened passively

Condition	AM1(Hz)	AM2 (Hz)
Simple AM	4	
	5	
	18	
	19	
	21	
	22	
Distant AM-AM	4	21
	5	22
Near AM-AM	18	21
	19	22



- Recording
  - Magnetic signals recorded using a 160-channel whole-head axial gradiometer system
  - Sampling rate 500 Hz, bandpassed between 1 Hz and 200 Hz, with notch at 60 Hz
  - 157 neural channels denoised with a Block-LMS adaptive filter, with 3 reference channels.

- Analysis
  - Discrete Fourier Transform (DFT) on the response; define aSSR: DFT's magnitude at modulation (*4,5, 18, 19, 21, 22 Hz*) and interaction (*3, 17, 25, 27, 39, 41Hz*) rates; 3 categories of interaction rates:

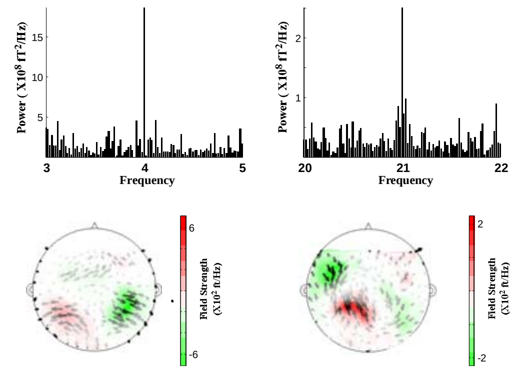
Condition	Distant AM-AM	Near AM-AM
Difference rate (Hz)	17 (21-4, 22-5)	3 (21-18, 22-19)
Sum rate (Hz)	25 (21 + 3), 27 (22 + 5)	39 (21 + 18), 41 (22 + 19)

  - Normalized aSSR: squared magnitude at the rate divided by the average squared magnitude of the spectral components ranging from 1 Hz below to 1 Hz above the target frequency, averaged over the 20 channels with the strongest normalized neural responses

Neural responses at modulation frequencies

- Power spectral density of MEG responses for a single subject, averaged over the 20 channels with highest signal-to-noise ratio (SNR). The stimulus is a broadband noise modulated by 4 Hz and 21 Hz concurrently.

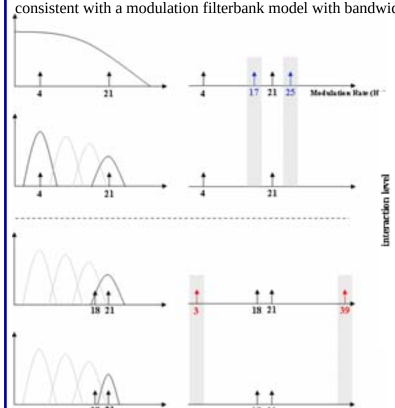
- The MEG magnetic field distributions of the neural response to modulations. The SSR at each channel is represented by an arrow whose length is proportional to the magnitude of the SSR and whose direction represents the phase. The head map is visually faded using the SNR of each channel as linear fading coefficients.

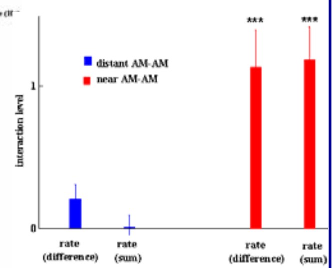


Neural responses at interaction frequencies

- Interaction Level: $IL = R_d - R_{nond}$

- Distant concurrent modulation rates do not evoke interaction responses, consistent with a modulation filterbank model over a single low pass modulation filter. Nearby concurrent modulation rates evoked strong interaction responses for associated sounds, relative with non-associated sounds, consistent with a modulation filterbank model with bandwidth greater than 3 Hz. .



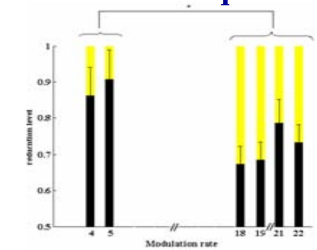


Responses change at modulation frequencies

- Reduction Level:

$$RL = \frac{R_s - R_c}{R_s}$$

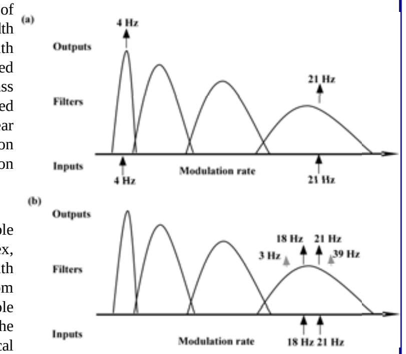
- The neural response is reduced for sounds with compound modulations. The reduction effect is more severe at higher rates. The star indicates a significant difference between low and high rates.



Discussion

- Schematic representations of modulation filter banks. The bandwidth of each filter is increased with increasing center frequency. The dashed line depicts the overall low-pass characteristic. Far AM-AM is resolved by two distinct filter banks while near AM-AM activate the same modulation filter bank, generating interaction components.

- The rate of the slowest detectable neural oscillation in auditory cortex, whether associated directly with modulation of stimuli, or resulting from interactions between multiple modulations also corresponds to the *perceived* rate of oscillation of physical stimuli.



Reference

[1] Shannon, R. V., Zeng, F.-G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech Recognition with Primarily Temporal Cues. Science, 270(5234), 303-304.

[2] Ross, B., Borgmann, C., Draganova, R., Roberts, L. E., & Pantev, C. (2000). A high precision magnetoencephalographic study of human auditory steady-state responses to amplitude-modulated tones. The Journal of the Acoustical Society of America, 108, 679.

[3] Draganova, R., Ross, B., Borgmann, C., & Pantev, C. (2002). Auditory cortical response patterns to multiple rhythms of AM sound. Ear Hear, 23(3), 254-265.

[4] Poeppel, D. (2003). The analysis of speech in different temporal integration windows: Cerebral lateralization as 'asymmetric sampling in time.' Speech Communication, 41(1), 245-255.

[5] Viemeister, N. F. (1979). Temporal modulation transfer functions based upon modulation thresholds. The Journal of the Acoustical Society of America, 66(5), 1364-1380

[6] Dau, T., Kollmeier, B., & Kohlrausch, A. (1997). Modeling auditory processing of amplitude modulation. II. Spectral and temporal integration. The Journal of the Acoustical Society of America, 102(5), 2906-2919.

Acknowledgements

Supported has been provided by NIH grants R01DC005660. We thank Mounya Elhilali for discussion and Jeff Walker for excellent technical support.