

TONE SEQUENCE STRUCTURE IS REFLECTED IN DYNAMIC NEURAL RESPONSES

Aniruddh D. Patel

The Neurosciences Institute

10640 John Jay Hopkins Drive, San Diego, CA 92121, USA

apatel@nsi.edu

ABSTRACT

Background. It is a commonplace notion that melodies are tone sequences which are neither too random nor too predictable in their structure. Little is known, however, about patterns of brain response as a function of the structure of tone sequences.

Aims. This study sought to determine if differences in the statistical structure of tone sequences are reflected in measurable, dynamic neural responses, and if sequences that are melody-like in their statistical properties have a distinct neural signature.

Methods. Subjects listened to 1-minute long diatonic tone sequences while neural signals were recorded using 148-channel whole-head magnetoencephalography (MEG). Sequences were random, deterministic (scalar), or one of two categories of 'fractal' sequences differing in their balance of predictability and unpredictability. (One of the fractal categories had melody-like statistics). Amplitude-modulation of the tone sequences was used to generate an ongoing, identifiable neural response whose amplitude and timing (phase) could be studied as a function of sequence structure.

Results. Ongoing timing patterns in the neural signal showed a strong dependency on the structure of the tone sequence. At certain sensor locations, timing patterns covaried with the pitch contour of the tone sequences, with increasingly accurate tracking as sequences became more predictable. In contrast, interactions between brain regions (as measured by temporal synchronization), particularly between left posterior regions and the rest of the brain, were greatest for the tone sequences with melody-like statistics. This may reflect the perceptual integration of local and global pitch patterns in melody-like sequences.

Conclusions. Dynamic neural responses reveal a neural correlate of pitch contour in the human brain, and show that interactions between brain regions are greatest when tone sequences have melody-like statistical properties.

1. INTRODUCTION

Melodies are a special subset of auditory sequences. Their acoustic raw materials can be extremely simple (e.g. a few dozen pure tones), yet the arrangement of these materials in time can create structures that engage a host of interacting mental processes, including chunking, melodic expectancy, and the perception of meter. Studying these processes is a principal goal for the cognitive science of melody, and is being actively pursued by a number of research groups (e.g. Krumhansl et al., 2000).

Discovering the neural correlates of melodic processing is a challenge for cognitive neuroscience. Progress in this area has focused on average neural responses to individual events in sequences (e.g. via event-related potentials or ERP Besson & Faïta, 1995) or on the brain's average response to entire sequences (e.g. via positron emission tomography or PET Zatorre et al., 1994). These techniques continue to provide valuable information, yet it is evident that to "tap into the moment-to-moment history of mental involvement with the music" (Sloboda, 1985), techniques are needed that measure patterns of neural activity as perception unfolds within individual sequences.

With these goals in mind, we set out to determine if aspects of tone sequence structure are reflected in dynamic neural responses, and if melody-like sequences have a distinct neural signature. Full details of this study are given in Patel & Balaban (2000). This paper emphasizes a qualitative understanding of our methods and results.

2. METHODS

To explore brain responses, we used statistically-generated tone sequences. This allowed us to generate novel stimuli which lay on a spectrum from random to deterministic in structure. We elected to use statistical tone sequences rather than precomposed melodies so that the sequences would be unfamiliar to subjects, easily generated in quantity, and mathematically well characterized. The

latter two points were of particular importance because we were employing a novel brain imaging technique and wanted to have good control over the stimuli.

All tone sequences were approximately 1 minute long, consisting of ~150 pure tones (415 msec each) with no temporal gaps. Sequences were diatonic, and ranged between A3 (220 Hz) and A5 (880 Hz) in pitch. Four structural categories of sequences were employed: random, deterministic (musical scales), and two intermediate 'fractal' categories of constrained variation which differed in their balance of predictability and unpredictability (Schmuckler & Gilden, 1993). These categories were given mathematical names in accordance with the technique used to generate them¹: $1/f$ ("one over f") and $1/f^2$ ("one over f squared").

A qualitative understanding of these categories is possible without delving into the underlying mathematics. In random sequences each successive pitch is chosen independently of the previous one, and there are no long term pitch trends. Deterministic sequences represent the opposite case: they consist entirely of long-term pitch trends (predictable stair-like patterns) with no short-term unpredictability. The fractal sequences are intermediate. $1/f$ sequences have a hint of long term pitch trends but still have much unpredictable variation from one pitch to the next. $1/f^2$ sequences are strong in long term pitch trends, but retain a small amount of unpredictability in the behavior of successive tones.

Examples of pitch contours from the different sequences are shown in Figure 1, f-i (black lines). Note the different shapes of the pitch contours in the four conditions: the random sequence has no discernable long-term patterns. The $1/f$ pitch contour has some evidence of long-term patterns (e.g. the general dip in the pitch contour in the middle of the sequence, followed by a slow climb in average pitch), but retains a good deal of unpredictable jagged pitch movement. The $1/f^2$ pitch contour has clearly discernable long-term pitch patterns with relatively little unpredictable jaggedness. The pitch contour of the scales moves up and down in a completely predictable way. Sound examples of all tone sequence categories can be heard at: www.nsi.edu/users/patel/tone_sequences.

Subjects (n=5 right handed males, 2 with musical training) were familiarized with the different stimulus categories in a training session where examples of each category were presented along with an arbitrary category label (the numbers 1-4). Subjects quickly learned to identify the different categories, and during the experiment, classified novel sequences by their category

with little difficulty. The experiment consisted of 28 such sequences, 7 per category. Stimuli in each category were equally distributed among seven Western diatonic modes (ionian, dorian, phrygian, lydian, mixolydian, aeolian, and locrian). Each subject heard a unique set of stimuli, with the exception of the scales, which were identical across subjects.

During stimulus presentation, neural data were recording using 148-channel whole head magnetoencephalography (MEG). MEG measures magnetic fields produced by electrical activity in the brain, providing a signal with similar time resolution to electroencephalography (EEG) but with certain advantages relating to source localization and independence of signals recorded from different parts of the sensor array (Lewine & Orrison, 1995).

We used a novel method to detect stimulus-related neural activity. Each sequence was given a constant rate of amplitude modulation (41.5 Hz), as shown in Figure 1 a-c. Fig 1a shows frequencies from a 4-second portion of a tone sequence. Fig 1b shows the associated amplitude waveform. Figure 1c provides a detail of a small piece of the waveform, showing the constant amplitude modulation frequency (41.5 Hz, blue line) overlaid on the changing carrier frequency. This amplitude modulation gave the tone sequences a slightly warbly quality, without disrupting their perceived pitch pattern: listeners heard them as sequences of pitches at the underlying pure tone frequencies.

It is known from auditory neuroscience that continuous amplitude modulation of pure tones results in a detectable brain response at the amplitude modulation frequency (Galambos et al, 1981; Hari et al., 1989), known as the auditory "steady-state response" (SSR). This response is visible in a power spectrum of the brain signal, which shows a peak at the amplitude modulation frequency. Fig 1d shows a 4-second piece of brain signal, and Fig 1e shows two corresponding power spectra, based on two successive 2-second portions of the signal. A peak at 41.5 Hz is clearly visible.

Thus amplitude modulation results in detectable stimulus-related cortical activity. We studied properties of this activity during individual sequences. In particular, for each sequence heard by a subject we studied the amplitude and timing characteristics (phase) of this activity in contiguous two-second epochs from each channel. One amplitude and phase value of the SSR was obtained from each successive 2-second epoch of the channel's brain signal via a Fourier transform, yielding approximately 30

data points x 148 channels per sequence.

Since a good deal of our analysis concerns phase information, it is worth giving a brief conceptual explanation of phase. By amplitude modulating our tone sequences at 41.5 Hz, we are introducing an oscillatory signal into the brain at that same frequency. This causes an oscillatory response (the SSR) at that frequency in certain brain regions. The degree to which the oscillatory brain response lagged the time-referenced input signal is measured by the phase of the brain response at 41.5 Hz. We studied the amplitude and phase of the brain response over time during individual sequences heard by our subjects.

3. RESULTS

Our first finding was that the phase of the measured brain signal varied with the pitch of the tone sequence. As pitch increased, phase advanced (corresponding to a *decreased* lag between stimulus brain response), and vice-versa. This general result is depicted in Fig 1e, which shows two spectra, one taken during a sequence of low pitches (Fig 1a, left half) and one taken when pitches were higher (Fig 1a, right half). Fig 1e shows that the peak of the SSR remains steady at 41.5 Hz, but its phase (inset arrow) advances as the average pitch of the tone sequence increases. This relationship between SSR phase and carrier frequency was suggested by early work (Galambos et al, 1981), and has been independently confirmed by another laboratory (John & Picton, 2000). It is likely to be due to the tonotopic layout of the basilar membrane in the human ear, where higher frequencies are closer to the oval window and hence stimulated earlier than lower frequencies.

Our next finding was that the phase of the brain response tracked the pitch contour a subject was hearing, and that this tracking improved as the sequences became more predictable in structure, with the best tracking for musical scales. Examples of phase-time contours (red lines) overlaid on their corresponding pitch time contours (black lines) are shown in Fig 1 f-i, showing how tracking improves across the stimulus conditions. Each subject showed a number of sensor locations where this 'phase tracking' of pitch was observed. Across subjects, these locations tended to be in fronto-temporal regions, with a right-hemisphere bias (Patel & Balaban, 2000, Fig 2). A similar set of locations was identified when we looked for sensors where the amplitude of the SSR was strong. However, we found no evidence that the amplitude of the SSR correlated with the heard pitch contour.

Knowing that the phase of the brain response contained information about stimulus properties, we then turned to looking at patterns of phase coherence between different brain regions. Phase coherence does not measure the lag between an oscillatory signal and brain response but rather the stability of the phase *difference* between oscillatory activity in different brain areas. Thus phase coherence is a measure of temporal synchronization between brain regions. If two brain areas show greater synchronization during certain condition, this is suggestive of a greater degree of functional coupling between those areas (see Bressler, 1995 for a review).

We found that across subjects, the different conditions were characterized by differing degrees of phase coherence. Random sequences generated less phase coherence than all other categories, and among the structured categories, $1/f^2$ sequences generated the greatest degree of phase coherence (Patel & Balaban 2000, Fig3). Interestingly, statistical research on Western music suggests that melodic tone sequences have approximately $1/f^2$ statistics (Nettheim, 1992; Boon & Decroly, 1995), suggesting that music-like sequences generated more brain interactions than other sequences.

To better understand the nature of these interactions, we examined topographic patterns of phase coherence, subdividing the brain into four quadrants (anterior and posterior x left and right). We found that the greater phase coherence of $1/f^2$ sequences was driven by interactions between the left posterior hemisphere and the rest of the brain, including the two right hemisphere quadrants. This is of interest because neuropsychological studies of brain-damaged patients suggests that left superior temporal regions are involved with the discrimination of precise interval sizes, while right fronto-temporal circuits are involved with the perception of more global contour patterns (Liégeois-Chauvel et al, 1998; Patel et al, 1998). Thus the observed pattern of coherence may reflect the dynamic integration of local and global pitch perception, and suggests that this integration is greatest when tone sequences resemble musical melodies.

4. DISCUSSION

This study has shown that it is possible to extract a signal from the human cerebral cortex which reflects the pitch contour an individual is hearing. The accuracy with which this signal reflects the pitch contour improves as the pitch sequence becomes more predictable. Thus there may be top-down influences of musical expectancy which influence this brain signal.

The basis of this signal is *temporal* information in cortical activity. When the *amount* of activity was examined, no relationship with pitch contour was observed. This suggests that dynamic imaging techniques have an important role to play in the study of music perception, complementing techniques sensitive to the amount of neural activity but insensitive to the fine temporal structure of that activity (e.g. functional magnetic resonance imaging, fMRI).

Dynamic imaging techniques also offer the opportunity to study how brain areas interact during perception. It is clear from decades of neural research that the brain is divided into different regions, each of which has a special role to play in perception and cognition. Yet it is also clear that these brain areas must interact to form coherent and unified percepts. Complex patterns such as music and speech engage multiple brain regions, and sequences with different perceptual properties may be distinguished by the pattern of brain interactions they engender rather than by the particular brain regions which respond to them.

Using phase coherence, we examined brain interactions as a function of stimulus structure and found that sequences with melody-like statistics engendered the greatest degree of neural interactions. In particular, we found evidence for strong functional coupling between the left posterior hemisphere and right hemisphere regions during the perception of melody-like sequences. This may reflect the perceptual integration of local and global pitch patterns, and suggests that one neural signature of melody is the dynamic integration of brain areas which process structure at different time scales.

Future work will use this technique to examine brain interactions as a function of stimulus structure in real melodies. This may provide one way to quantify the perceptual coherence of melodies in individuals who cannot easily give details of their perception, such as non-musicians and infants.

ACKNOWLEDGEMENTS

This work was supported by the Neurosciences Research Foundation as part of its research program on Music and the Brain at The Neurosciences Institute.

NOTES

1. Inverse Fourier transform of power spectra with different slopes (see Patel & Balaban 2000 for details).

REFERENCES

- Besson, M. & Fäita, F. (1995). An event-related potential (ERP) study of musical expectancy: Comparison of musicians with nonmusicians. *J. Exp. Psych: Human Perception and Performance*, 21, 1278-1296.
- Boon, J.P. & Decroly, O. (1995). Dynamical systems theory for music dynamics. *Chaos* 5, 501-508.
- Bressler, S. (1995). Large-scale cortical networks and cognition. *Brain Research: Brain Research Reviews*, 20(3), 288-304.
- Galambos, R., Makeig, S. & Talmachoff, P.J. (1981). A 40-Hz auditory potential recorded from the human scalp. *Proc. Natl. Acad. Sci. USA* 78, 2463-2647.
- Hari, R. Hämäläinen, M., & Joutsiniemi, S.-L. (1989). Neuromagnetic steady-state responses to auditory stimuli. *J. Acous. Soc. Am.* 86, 1033-1039.
- John, M.S. & Picton, T.W. (2000). Human auditory steady-state responses to amplitude modulated tones: phase and latency measurements. *Hearing Research*, 14, 57-79.
- Krumhansl, C., Louhivuori, J., Toiviainen, P., Järvinen, T. & Eerola, T. (2000). Melodic expectation in Finnish spiritual folk hymns: convergence of statistical, behavioral, and computational approaches. *Music Perception*, 17(2), 151-196.
- Lewine, J.D. & Orrison, W.W. (1995). Magnetoencephalography and magnetic source imaging. In: *Functional Brain Imaging* (W.W. Orrison et al., ed): 369-417. St. Louis: Mosby.
- Liégeois-Chauvel, C., Peretz, I., Babai, M., Laguitton, V. & Chauvel, P. (1998). Contribution of different cortical areas in the temporal lobes to music processing. *Brain* 121, 1853-1867.
- Nettheim, N. (1992). On the spectral analysis of melody. *Interface* 21, 135-148.
- Patel, A.D. & Balaban, E. (2000). Temporal patterns of human cortical activity reflect tone sequence structure. *Nature*, 404, 80-84.
- Patel, A.D., Peretz, I., Tramo, M. & Labrecque, R. (1998). Processing prosodic and musical patterns: a neuropsychological investigation. *Brain and Language* 61, 123-144.
- Schmuckler, M.A. & Gilden, D.L. (1993). Auditory perception of fractal contours. *J. Exp. Psychol: Human Percep. & Perform.* 19, 641-660.
- Sloboda, J. (1985). *The Musical Mind: The Cognitive Psychology of Music*. Oxford: Clarendon Press.
- Zatorre, R.J., Evans A.C. & Meyer, E. (1994). Neural mechanisms underlying melodic perception and memory for pitch. *Journal of Neuroscience*, 14(4), 1908-1919.

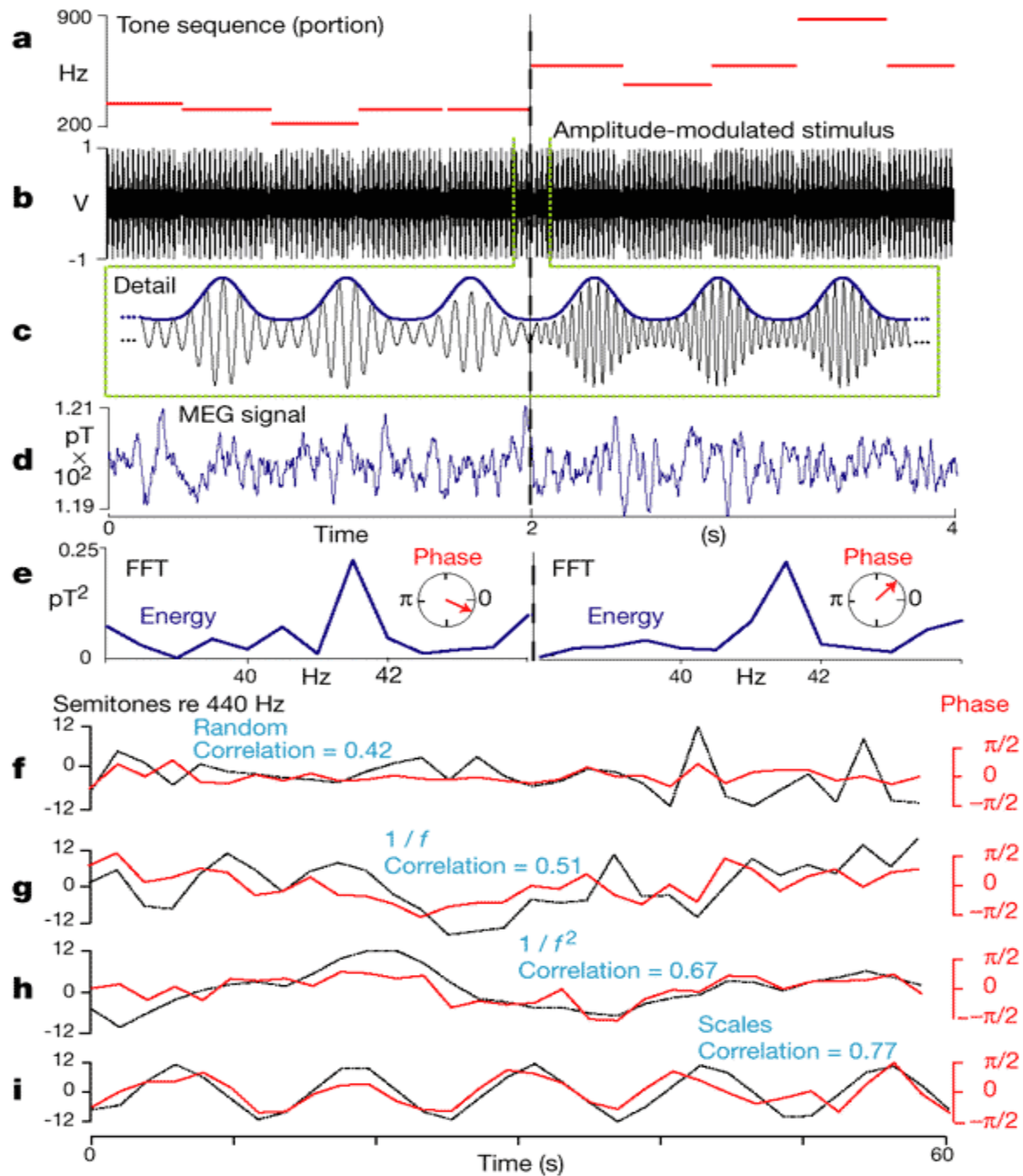


Figure 1. (a-d): Example of stimulus and brain response over 4 seconds: (a) Tone frequencies; (b) Stimulus waveform; (c) Waveform detail (150 msec), showing constant modulating frequency (41.5 Hz, blue line) overlaid on changing carrier frequency; (d) Neural signal from one sensor. (e) Successive 2-second spectra of neural signal. The brain signal shows an energy peak at 41.5 Hz, whose phase (inset arrow) varies with carrier frequency. (f-i): Phase-tracking of individual tone sequences. Pitch-time contours (black lines) illustrate the four different stimulus categories. Associated neuromagnetic phase-time series (red lines) from a single sensor during one trial in one subject were randomly drawn from the top 10% of sensor correlation values for each stimulus. The correlation between the resampled pitch-time series and the neuromagnetic phase-time series is given in the inset to each graph.