Adolfo G. Ramirez-Aristizabal

Experiment 1 Report

EEG Response and Auditory Stimuli Complexity Matching

Abstract

Coordination of cognitive systems have been seen to follow complexity matching principles in their information exchange. Experimental approaches have sought methods to quantify the statistical complexity of signals in these systems to infer the extent of matching that may occur during interaction. One class of coordination demonstrates a system that plays the role of the perturber and another one as the perturbed. In the present study, auditory stimuli are presented as perturbations to auditory processing in the brain, and we test whether these perturbations can be measured in EEG signals through complexity matching. A new approach is used to try to capture hierarchical temporal structure of EEG recordings from natural auditory stimuli that allow for longer, natural stimuli lengths without the limits of repeated stimuli exposure. Current results of this methodology affirm the potential for capturing hierarchical temporal structure in EEG responses but complexity matching stays elusive through limitations of hierarchical processing localization.

Introduction

Music and language are both constructed from smaller to larger units that hierarchically organize content e.g., determiner & noun to build a noun phrase in language and varying note pitches that build a larger harmonic phrase. This hierarchical construct has been predicted to be parsed by similar neural pathways whether the incoming information is categorized as music, or language (Patel, 2003). Let us look at auditory perception research starting from simple stimuli responses to complex multilevel responses, to build up to hierarchical response processing.

Neural entrainment research has demonstrated that neural oscillations are sensitive to temporal features, as seen when participants synchronize to perceived metrical onsets of an isochronous auditory stimuli (Tal et al., 2017). Furthermore, faster metrical rhythms show a corresponding array of entrained neural responses that are also oscillating at faster frequencies as well as slower responses for slower rhythms (Musacchia, Large, & Schroeder, 2014). In addition, pitch variance in auditory stimuli can be used to induce a stronger meter percept. Stronger identification has been shown for stimuli that have shorter semitone distances, which further tests the multidimensional aspect of complex auditory stimuli (Jones et al., 2006).

With respect to complex stimuli analyzed hierarchically, the use of natural speech stimuli presented at isochronous intervals has been used to demonstrate hierarchical cortical tracking of semantic units (Ding, Melloni, Zhang, Tian, & Poeppel, 2015). Neural responses show tracking occurring at the sentence level (1Hz), phrase level (2Hz), and syllable level (4Hz). Similarly, evidence for a hierarchical system in auditory processing is posited by demonstrating tracking at lower frequencies of phonemic units in natural speech ~155s (Di Liberto, O'Sullivan, & Lalor,

2015). Longer stimuli are also used in the attempt to analyze hierarchical processing of music and speech with lengths closer to their natural lengths ~4:15 mins (Farbood, Heeger, Marcus, Hasson, & Lerner, 2015). They present conditions with scrambled structure at three different timescales of the music separately to trained musicians. Using fMRI, larger musical timescales were found to be processed longer when approaching higher order brain topography in the auditory cortex.

Given that human complex networks are highly adaptable (Baronchelli, Ferrer-i-Cancho, Pastor-Satorras, Chater, & Christiansen, 2013), we investigate the hypothesis that the dynamics of brain responses to complex acoustic signals may be systematically related to the dynamics of the signals themselves. Consistent with our approach, Lalor & Foxe (2010) use Auditory Evoked Spread Spectrum Analysis to extract responses with high temporal resolution over continuous auditory stimuli e.g., they use a 48 min speech recording cut up into 16 blocks. Their stimuli driven method uses least squares estimation to clean comodulating differences between input and output for a cleaner signal, and location for EEG recording comes from select channels such as Fz. Skoe & Kraus (2010) describe a method for finding neural synchrony to auditory stimuli using complex auditory brainstem responses (cABR). Stimuli take a short clip of any complex auditory stimuli, typically no more than 10 seconds long, and present it ~2000-3000 times. The collected cABRs have high temporal and spectral resolution which holds enough information to reconstruct the signal and play it back, sounding very similar to the original stimulus. Subcortical regions can also have a role in the temporal selection of a beat during entrainment, but the critical shaping occurs cortically with stronger effects for complex rhythms (Nozaradan, Schönwiesner, Keller, Lenc, & Lehmann, 2018). Both speech and music have shared activations in the auditory cortex (Patel, 2003), in which hierarchical processing can be seen to occur. Differences of natural speech not being isochronous or having a beat hints at corresponding functional activation occurring in ventrolateral regions (Farbood et al., 2015). Hierarchical linguistic units of speech have been tested and found to be more specifically related to temporal lobe regions such as the superior temporal gyrus and the superior temporal sulcus (Di Liberto et al., 2015; Ding et al., 2015).

Methods to quantify hierarchical temporal structure (HTS) have been used by (Kello, Bella, Médé, & Balasubramaniam, 2017) on an array of natural acoustic signals including varying genres of music, conversations, speeches, and animal vocalizations. The HTS of each acoustic signal was calculated by the Allan Factor (AF) statistic, which gives a score of variance of peak amplitude events at a given time scale. Given a set of time scales it produces a function of how variability of events scales up, to describe the nesting relationship of the signal. Thunderstorms and classical music showed similar HTS as well as Jazz and conversations despite their perceived differences. Embellishment in speech as in the case of infant directed speech compared to adult directed speech is seen in the HTS of the signal as an embellishment at the longer timescales (Falk & Kello, 2017). The HTS of acoustic signals has also been shown to be affected by speaking rate and naturalness; showing a decrease in HTS at the longer timescales for synthesized speech over natural speech and a compression of HTS to the shorter timescales at faster speaking rates (Ramirez-Aristizabal, Médé, & Kello, 2018).

Given a method that can analyze HTS of acoustic signals, the present study aims to investigate whether this same method can capture HTS in a continuous brain response to

complex acoustic stimuli captured through EEG. The extent to which the brain response is related to the HTS of the acoustic stimuli is tested through principles of complexity matching. First discussed by West, Geneston & Grigolini (2008), the Complexity Matching Effect states that maximal information transfer between complex networks happen when their temporal complexities match. This is measured by looking at the exponent of their $1/f^{\mu}$ distributions from the inter-event-intervals of crucial events in the signals produced by each network ($\mu_1 = \mu_2$). Behavioral experiments first tested this concept through dynamics of finger tapping and pendula being swung together (Stephen, Stepp, Dixon, & Turvey, 2008; Marmelat, & Delignières, 2012). Fractal metronome input were followed closely by participants and the fluctuations of their inter-tap intervals matched the power law exponent of the metronome. Similarly, angular fluctuations between two participants swinging a pendulum followed a $1/f$ distribution but the interactions were not fixed as in the metronome example because the dynamic locations of the pendulums affected the other causing a bidirectional perturbation. An adaptation of complexity matching going away from matching μ exponents was seen in a study by (Abney et al., 2014) using the Allan Factor method. Conversations between dyads were analyzed by looking at the distributions of their inter-onset-intervals. Instead of equating the distribution exponents, the power law clustering of events that reflect the hierarchical temporal structure of language was used to find complexity matching occurring in affiliative conversations.

The goal of this study is to test the Allan Factor method to capture HTS in EEG responses and investigate the extent that complexity matching principles can be applied. Stimuli are taken from the Kello et al., (2017) study to test an array of distinct types of sounds and unique HTS categories. Each type of stimuli is only presented once at full length of recording. Time series distributions are parsed from peak amplitude events in the EEG recording. Independent component analysis is used to localize data in topographic points of interest as discussed above within frequency ranges of 1 - 50 Hz. Components located in the shared regions of activation (auditory cortex) will be the primary focus for testing HTS preservation in brain responses.

Methods

Data Collection

Our approach sought to collect cortical responses through EEG to an array of 6 auditory stimuli e.g., electronic music, bird song, ted talk, sine transform of ted talk, classical music, and the same classical song repeated. The lengths of the stimuli ranged from 4:20 – 4:42 mins and the EEG recording lengths where 4:20mins. The audio was down sampled to the sampling rate of the EEG system (32 channel ANT Waveguard electrode cap) at 2056 Hz. There was a total of 11 participants and the presentation of the stimuli was randomized between subjects. Headphones were used and kept at comfortable levels for the participants. During the presentation of stimuli, participants were told to stare at a black screen. Collected data was further processed to remove artifacts and unwanted noise. The removal of 60 Hz sinusoids was implemented to clean up noise from electrical appliances. A 0.1 Hz high pass filter was used to remove drift from data. Bad channels were rejected using a probability function.

ICA

Adaptive Mixture Independent Component Analysis (AMICA) was used to localize EEG components that topographically cluster in the auditory cortex and other points of interest. ICA weights were retrieved from modified recordings which were down sampled from 2048 Hz to 1024 Hz and band passed at 1-50 Hz. Then those weights were transferred to the original EEG recordings but keeping it down sampled at 1024 Hz. The EEGLab plugin AMICA was used with 3 models using posterior probabilities for each model to keep the best components among the different models. K-means clustering was used taking into account spectral, and topographic information of the components across participants as seen in Figure 1. An n of 9 was used, which placed 9 distinct means and classified components with similar means into a cluster. The n parameter was also a practical choice, because it allowed for the creation of a cluster near the auditory cortex with at least one component per participant and condition. The Allan Factor of the auditory cortex cluster components were taken to test HTS.
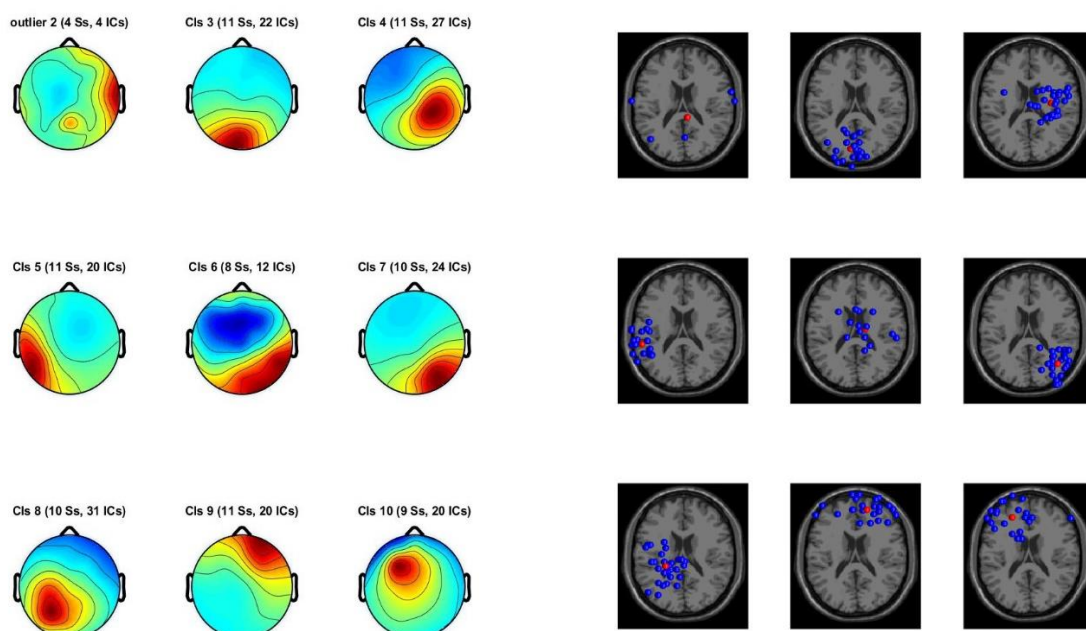

Figure 1: AMICA components of individual participants where put through K-means clustering and broken down into 9 clusters containing components across participants. First panel (left) shows a heatmap of spectral clustering and the second panel (right) breaks down components by dipole fitting locations. The first cluster at the top left of both panels contains components that were located outside the cortex and treated as artifacts, as well as components that were over 15% variance in the group.

Allan Factor

Auditory stimuli and EEG components were analyzed using Allan Factor. First, signals were divided into four-minute segments, and the Hilbert envelope was calculated for each segment. Analyses were averaged across segments to yield a single AF function per signal. Peak amplitude events were then selected using two parameters. The first parameter identified maximal peaks within a ±5 ms window, then all peaks were preserved if their amplitudes were

over the *H* parameter, a breakdown of this process is seen in Figure 2. Selected events created a time series in which the AF statistic was implemented on as follows:

$$A(T) = \frac{\left\langle \left(N_i(T) - N_{i+1}(T)\right)^2 \right\rangle}{2\langle N_i(T) \rangle}$$

where *T* is the timescale, $N_i(T)$ is the event count in each window *i*, and $A(T)$ is AF variance. AF variance captures the degree of event clustering at a given timescale, and for a time series with nested clustering, $A(T)$ increases with *T*. Self-similar clustering across timescales yields a power law, $A(T) \sim T^{\alpha}$, where $0 < \alpha < 2$. The AF function was computed for 11 values of *T* in between 15 ms and 15 s, logarithmically spaced to compute the orthonormal basis.
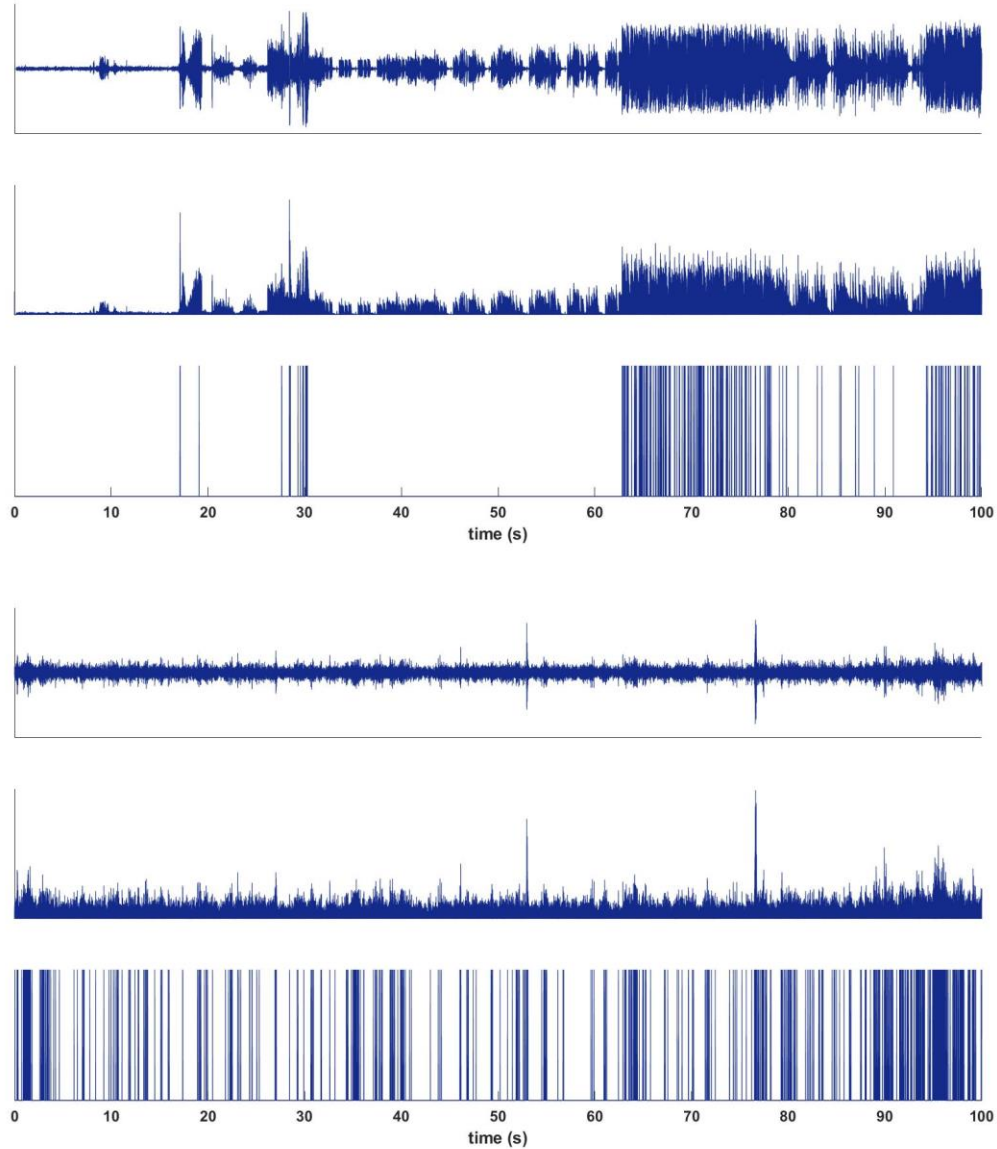

Figure 2: An example of peak amplitude event selection into a time series is shown for the electronic dance music stimulus (top) and a corresponding EEG response (bottom) of a 100 second length. For both top and bottom panels, the first section shows the signal waveform, followed by it's Hilbert envelope, and lastly the time series in which the AF statistic is applied to.

Classification

Machine learning through Matlab's Classification Learning application was used to test for minute variations not captured in our inferential statistics. A matrix of the data was formatted using the AF scores of each ICA component from every participant and from the four cluster. The targeted response for classification was the six stimuli categories and the predictors were the eleven AF values which account for the eleven timescales in the analysis. Different combinations of predictors were tested as well by adding the cluster labels and another case which tested the response variable of the four clusters for classification. Training for classification used all the Support Vector Machine (SVM) options which include linear, quadratic, cubic, fine gaussian, medium gaussian, and coarse gaussian SVMs. This approach for using AF values for label classification is similar to the approach by Kello et al., (2017).

Results

The first set of results presented demonstrate clear distinctions in AF function shapes for the stimuli (Figure 3). An exception is observed in the similarity of event clustering at the longer timescales for the Ted talk and its sine wave transform.
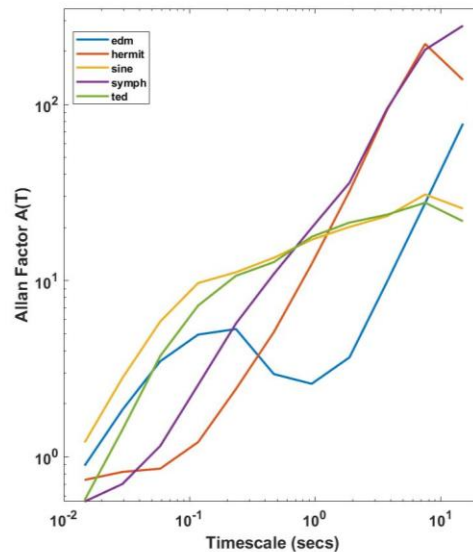

Figure 3: AF functions for the down-sampled auditory stimuli are presented above.

The next set of analyses took ICA decompositions of individuals and organized them into 9 clusters through K-means in order to locate points of interests of brain responses in the cortex. The initial focus was on doing AF on ICA components that fell into a cluster close to the auditory cortex, but the shapes of AF functions looked similar to $1/f$ scaling. We then tested to see if this persisted across clusters. We only took into account clusters that had at least one component for each participant per condition. Four clusters met that criteria and they belonged to the frontal cortex, occipital, sensory-motor (right hemisphere), and auditory cortex (left

hemisphere) regions (Figure 1). All clusters were shown to follow the same $1/f$ pattern despite their differences on topography. The clusters were averaged together and plotted by condition (Figure 4). Despite AF functions having $1/f$ scaling throughout, the longest timescales showed a notable difference. Linear fit slopes of the AF functions of all components where taken to test for differences. Figure 4 shows the mean linear slopes of ICA components when organized by condition and by order of condition presented. A two-way repeated measures ANOVA of condition linear slopes shows a p-value of 0.0633 and the one for condition order has a p-value of 0.0969. A Tukey's Honest Significant Difference Test is performed on both conditions' and condition orders' linear slopes with the only significant difference occurring in between the Symphony 1 and Symphony 2 condition slopes with a p-value of 0.0471.

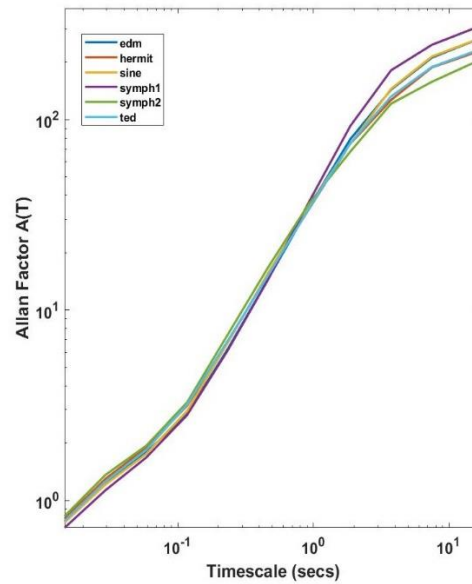|	EDM	Hermit	Sine	Symph1	Symph2	Ted
Condition Means	1.0457	0.9587	0.9858	1.1355	0.9237	0.9799
Condition Standard Errors	0.0504	0.0565	0.0543	0.0575	0.0462	0.0495
	1	**2**	**3**	**4**	**5**	**6**
Order Means	0.9664	0.999	1.0815	1.0815	1.0154	0.8928
Order Standard Errors	0.0537	0.0509	0.0615	0.0519	0.0542	0.0439

Figure 4: Aggregated Allan Factor functions of ICA components are averaged by condition across clusters (top). The average AF slopes from ICA components organized conditions and condition order (bottom) and their

standard errors are presented. Slopes come from a linear fit on the longest timescales where the biggest differences exist.

Classification learning of AF values demonstrated results that support the above presented statistics. The first case had only AF values as predictors and the stimuli categories as response variables. In all SVMs the classification did not go above chance 16.67%. When we added cluster labels and participant labels as predictors it would lower classification performance. We also tested the shorter timescales and the longer timescales by breaking up the classification testing in two. The shorter timescales used only the first three AF values as the predictors and yielded lower performance than chance, with the best performing SVM yielding 15.7%. The longer timescales used the last three AF values and had performance higher than chance at 22.3% with a fine gaussian SVM. The second classification case involved the cluster labels as response variables and the AF values as predictors. The best SVM classification came from the quadratic SVM with a performance of 28.4% above chance from the four cluster categories. The frontal lobe cluster showed to be the most unique in its classification performance with the highest rate of true positives at 68%.

Discussion

The aim of this study was to test for complexity matching between distinct auditory stimuli and their corresponding brain responses using EEG. The stimuli used are in reference to Kello et al., (2017), in which the stimuli with the most distinct AF functions were used as well as a second presentation of the symphony stimulus to test effects of repetition. The AF functions in Figure 3 show a replication of the characteristic AF functions that electronic music, Hermit Thrush bird song, Ted talk, Symphony, and the Sine transform of the Ted talk. ICA components did not demonstrate AF functions with similar corresponding event cluster scaling.

The AF functions of brain responses at the individual EEG channel level and ICA compositions from different topographic cluster locations all yielded AF functions akin to $1/f$ event clustering. This finding replicates studies of spontaneous brain activity of healthy brains which is explained as a sign of self-organized criticality (Freeman, 2005). Complexity matching principles were tested by Allegrini et al., (2009) as $1/f$ brain dynamics are shown to have optimal $\mu = 2$ scaling. They propose that the brain should be most sensitive to $1/f$ signals as seen in classical music and other phenomena in nature. Our results did not support this proposal, because HTS of brain responses were found to be $1/f$ despite the original stimuli not necessarily being $1/f$. This result held for all clusters including the ICA components near the auditory cortex.

We should note a few limitations of the present experiment that qualify our conclusions and warrant further investigation. First, although stimulus order did not come out significant (p = 0.0969), orders were not balanced across participants. Next, a larger sample size would be helpful in clarifying the two-way ANOVA statistic between AF slopes and types of stimuli used.

Finally, a larger array of EEG electrodes would help with spatial resolution and possible isolation of relevant brain responses.

One mediating factor affecting brain responses is neural modulation due to familiarity and prediction errors (Chang, Bosnyak, & Trainor, 2016). Specifically, beta band power shows an increase in entrainment trials when an induced prediction error occurs. Dynamic Attending theory states that the brain's endogenous self-sustained neural oscillations and the dynamic allocation of attention are parameters that would directly influence modulation in a brain response (Jones & Boltz, 1989; Large & Jones, 1999; Jones, 2010). In other words, self-sustained oscillations can be thought of as a "default mode" in which the presentation of auditory stimuli affects the rate and phase of the oscillations. This approach leads one to consider neural modulation due to novelty of stimuli that might result in higher variability in the beta frequency band, and more sensitivity to prediction errors. Future studies may clarify the role of these and other factors, perhaps by examining entrainment specifically. For instance, stimuli could be repeated several times and the AF slopes of the repeated stimuli would be compared for significant change due to entrainment/familiarity. Our results hint at this effect already because the biggest difference in AF slopes happened between the first presentation of the symphony stimuli and the second presentation ($p = 0.0471$). The classification learning also demonstrated that most of the differences that would allow an SVM to classify signals properly, were happening in the longer timescales. Meaning that the data does hold information about differences in stimuli presentation and should be further tested in the above-mentioned experimental paradigm.

Another task for future work is to explore ways to measure HTS of brain responses for stimuli complexity matching, such that measures are more directly related to the stimulus effects. Our attempt at separating the signal by targeting specific ICA components in related topographic regions of the brain did not demonstrate AF functions with obvious relations to the stimuli. HTS of ICA components clustering around the auditory cortex were surprisingly similar to those components in the occipital, sensory-motor, and frontal regions. However, hierarchical processing of both music and speech have been found to yield different patterns of activation topographically along the A1 auditory cortex, up to the temporal-parietal junction (Lerner, Honey, Silbert, & Hasson, 2011; Farbood et al., 2015). The classification learning procedure also showed that with our data, the frontal lobe was the most unique cluster of components. Future studies may employ similar methods of analysis to test whether hierarchical processing in these brain regions is affected by HTS in auditory stimuli. Such studies would benefit from a timescale-localization model for auditory processing to better guide a cleaner extraction of the corresponding brain response to presented stimuli.

References

Abney, D. H., Paxton, A., Dale, R., Kello, C. T., Abney, D. H., Paxton, A., … Kello, C. T. (2014). Journal of Experimental Psychology : General Complexity Matching in Dyadic Conversation Complexity Matching in Dyadic Conversation.

Allegrini, P., Menicucci, D., Bedini, R., Fronzoni, L., Gemignani, A., Grigolini, P., … Paradisi, P. (2009). Spontaneous brain activity as a source of ideal 1/f noise. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, *80*(6), 1–13. https://doi.org/10.1103/PhysRevE.80.061914

Baronchelli, A., Ferrer-i-Cancho, R., Pastor-Satorras, R., Chater, N., & Christiansen, M. H. (2013). Networks in Cognitive Science, *17*(7), 348–360. https://doi.org/10.1016/j.tics.2013.04.010

Chang, A., Bosnyak, D. J., & Trainor, L. J. (2016). Unpredicted pitch modulates beta oscillatory power during rhythmic entrainment to a tone sequence. *Frontiers in Psychology*, *7*(MAR), 1–13. https://doi.org/10.3389/fpsyg.2016.00327

Di Liberto, G. M., O'Sullivan, J. A., & Lalor, E. C. (2015). Low-frequency cortical entrainment to speech reflects phoneme-level processing. *Current Biology*, *25*(19), 2457–2465. https://doi.org/10.1016/j.cub.2015.08.030

Ding, N., Melloni, L., Zhang, H., Tian, X., & Poeppel, D. (2015). Cortical tracking of hierarchical linguistic structures in connected speech. *Nature Neuroscience*, *19*(1), 158–164. https://doi.org/10.1038/nn.4186

Farbood, M. M., Heeger, D. J., Marcus, G., Hasson, U., & Lerner, Y. (2015). The neural processing of hierarchical structure in music and speech at different timescales. *Frontiers in Neuroscience*, *9*(APR), 1–13. https://doi.org/10.3389/fnins.2015.00157

Freeman, W. J. (2005). A field-theoretic approach to understanding scale-free neocortical dynamics. *Biological Cybernetics*, *92*(6), 350–359. https://doi.org/10.1007/s00422-005-0563-1

Jones, M. R., & Boltz, M. (1989). Dynamic Attending and Responses to Time. *Psychological Review*, *96*(3), 459–491. https://doi.org/10.1037/0033-295X.96.3.459

Kello, C. T., Bella, S. D., Médé, B., & Balasubramaniam, R. (2017). Hierarchical temporal structure in music, speech and animal vocalizations: Jazz is like a conversation, humpbacks sing like hermit thrushes. *Journal of the Royal Society Interface*, *14*(135). https://doi.org/10.1098/rsif.2017.0231

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological review*, *106*(1), 119.

Lerner, Y., Honey, C. J., Silbert, L. J., & Hasson, U. (2011). Topographic Mapping of a Hierarchy of Temporal Receptive Windows Using a Narrated Story. *Journal of Neuroscience*, *31*(8), 2906–2915. https://doi.org/10.1523/JNEUROSCI.3684-10.2011

Marmelat, V., & Delignières, D. (2012). Strong anticipation: complexity matching in interpersonal coordination. *Experimental Brain Research*, *222*(1), 137–148.

https://doi.org/10.1007/s00221-012-3202-9

Ramirez-Aristizabal, A. G., Médé, B., & Kello, C. T. (2018). Complexity matching in speech: Effects of speaking rate and naturalness. *Chaos, Solitons & Fractals*, *111*, 175–179. https://doi.org/10.1016/j.chaos.2018.04.021

Stephen, D. G., Stepp, N., Dixon, J. A., & Turvey, M. T. (2008). Strong anticipation: Sensitivity to long-range correlations in synchronization behavior. *Physica A: Statistical Mechanics and Its Applications*, *387*(21), 5271–5278. https://doi.org/https://doi.org/10.1016/j.physa.2008.05.015

Tal, I., Large, E. W., Rabinovitch, E., Wei, Y., Schroeder, C. E., Poeppel, D., & Zion Golumbic, E. (2017). Neural Entrainment to the Beat: The "Missing-Pulse" Phenomenon. *The Journal of Neuroscience*, *37*(26), 6331–6341. https://doi.org/10.1523/JNEUROSCI.2500-16.2017