Functional Connectivity and information flow in ECoG data: A Linear Approach.

Tim Mullen

STAT 153 (TSA) / Prof. Brillinger / Spring 2007

1. INTRODUCTION AND BACKGROUND

An interesting problem in neuroscience is determining the functional connectivity and information transfer properties of neuronal networks engaged in cognitive information processing. For example, we might want to know whether there is a causal relationship between two cortical areas, and, if so, at what temporal delay does information propagate. Although it is widely accepted that neuronal processes exhibit non-stationary, non-linear behavior, previous work has shown that linear time series analysis methods such as cross-correlation and coherence may be suitable for exploring the interaction between these processes. In previous work [1], I have explored the use of mutual information and cross-correlation in identifying information propagation latencies between several cortical regions during language comprehension and production tasks. However, in that work I did not compensate for the non-stationarity of the data or for the effect of high autocorrelations on cross-correlation estimates. In this paper I will explore the use of differencing and autoregressive prewhitening prior to application of linear system identification techniques, such as cross-correlation and coherence in the hopes of obtaining a refined estimate of the interactions between language, auditory, and motor processing areas while a language comprehension task is being performed. All analyses and figures in this paper were computed in MATLAB.

The motor theory of language [2] proposes that language was modeled on preexisting neural systems responsible for motor control; complex interaction between these elementary motor programs formed the basis for language development. Phonetic structure is a result of the prior organization of motor programs which enervated various portions of the articulatory system. Phonemic perception then takes place by activation of intrinsically-linked sensorimotor subprograms while higher-level interaction between groups of these subprograms gives rise to language (e.g., word, sentence) comprehension. This

suggests that language processing is facilitated by means of a complex spatially-distributed network linking sensorimotor cortex along with more traditionally accepted auditory and language processing areas.

Recent studies [3, 4] have provided evidence that language/auditory and motor/pre-motor areas are functionally connected within a network with estimated propagation delays of approximately 10-36 ms (estimated from Magnetoecephalograph (MEG) amplitude peak-peak estimates and direct cortico-cortical stimulation). Hence, if the motor theory is correct, a language processing task should engage this network and we may expect to find directionally-specific information flow between auditory and motor regions. Although this interaction may be nonlinear, nonlinear system identification can be rather challenging to carry out in practice. A simplified approach, which may provide an adequate approximation to the true interaction, is to assume a linear, time-invariant system whose information transfer properties we may analyze using classical linear systems analysis. Prior analysis of electrocorticogram (ECoG) data collected during a language comprehension task showed that time-delayed mutual information (which takes into account nonlinear dependencies) and cross-correlation yielded significantly similar estimates of peak latencies and connectivity strength, bolstering the suitability of a linear approach. This paper therefore focuses on analysis of directionally-specific information flow between primary auditory (A1) and primary motor cortex (M1) in a single patient during a language comprehension task. The two regions analyzed are shown in Figure 1.

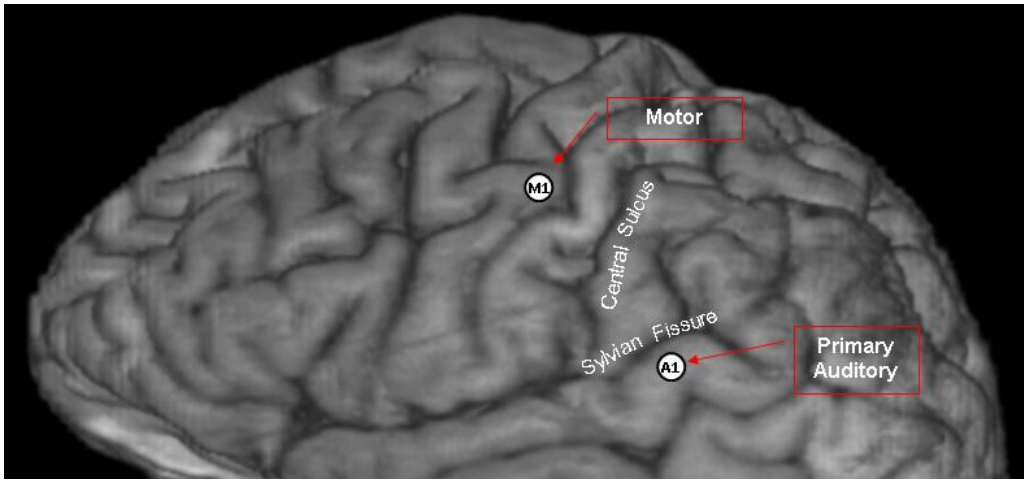

Figure 1. Anatomical MRI of Patient 1 showing locations of channels analyzed.

2. METHODS

Multi-channel subdural electrocorticogram (ECoG) data was recorded from patients undergoing neurosurgery for epilepsy. The 64-channel grid is implanted for 4-7 days, during which experimental recording is performed. This paper focuses on data from a single patient performing a simple language comprehension task. Subjects are presented aurally with words (verbs), non-words (scrambled phonemes with same spectral properties as real words), and proper nouns and instructed to press a button when they hear a proper noun. ECoG is collected at a sampling rate of 2003 Hz. Since an external reference electrode was not available, each electrode is clinically referenced to a grid electrode (#64), and later re-referenced to a common-average. Electrodes with poor signal-to-noise ratio were removed prior to common-averaging. To improve autoregressive model fitting capability, the data was downsampled to 1024 Hz. A zero-phase digital FIR filter with a 3-160 Hz passband was then applied to remove DC drift, residual ocular movement artifacts (they are vastly reduced in ECoG), as well as a strong 180 Hz line noise harmonic. For each condition {word, non-word, proper-noun} each channel's time-series was segmented into 96 non-overlapping epochs each representing the interval 200 ms pre-stimulus to 800 ms post-stimulus for a particular stimulus presentation. Proper noun segments were discarded since they contain button-press motor artifacts. For this paper I chose to examine only the {word} stimulus.

3. ANALYSIS

3.1. The Linear Delay Model

One approach to linear system identification is to assume we have an open-loop system of the form

$$(1) \qquad Y_t = \sum_{k=0}^{\infty} h_k X_{t-k} + N_t$$

with "input" X_t, "output" Y_t, impulse response function $\{h_k\}$, and uncorrelated white noise N_t. We are primarily interested in identifying the delay of the system, given by the shape of the impulse response function. Two tools we will make extensive use of are the autocorrelation (a.c.f.) and cross-correlation (c.c.f.) functions. Assuming that X_t and Y_t are stationary zero-mean processes, these are given

by $\rho_X(k) = \mathrm{E}[X_t, X_{t+k}]$ and $\rho_{XY}(k) = \mathrm{E}[X_t, Y_{t+k}]$, respectively, where $\mathrm{E}[\cdot]$ is the expected value. If one multiplies both sides of equation 1 by X_{t-q} and takes expectations, it can be shown that, unless $\{h_k\}$ is uniformly zero, $\rho_{X->Y}(q)$ depends on $\rho_X(q-k)$ and it becomes difficult both to interpret the c.c.f. and to estimate $\{h_k\}$. If, however, the input and output are suitably transformed such that the input series is significantly autocorrelated only at lag zero, then we find that $\{\rho_{XY}(0), \rho_{XY}(1), ..., \rho_{XY}(k)\} = \{h_k\}$. In other words, the c.c.f. gives us precisely the impulse response function of the system (Chatfield, 196). The Box-Jenkins method of model identification (Chatfield section 9.4.2) utilizes a method known as autoregressive prewhitening to de-correlate the input series: we fit an ARMA/ARIMA model to X_t and obtain the residual series, X_t'. If a good fit was found, X_t' should be approximately uncorrelated white noise. We then filter Y_t using the same AR model and obtain the residual, Y_t'. Cross-correlations may then be carried out on the residuals to determine the delay of the system. Coherence, phase, and gain estimates may also benefit from the prewhitening (Brillinger, personal correspondence). Box and Jenkins further utilize $\{h_k\}$ to estimate additional parameters of a more parsimonious version of equation 1. However, in this paper I was primarily concerned with obtaining an accurate estimate of the functional connectivity and temporal delay based on cross-correlation and coherence between the two series, unencumbered by the problems associated with high autocorrelations and non-stationarity. Thus I chose to defer full estimation of the model and instead focus on these initial goals.

3.2. Preprocessing

I utilized the following procedure to compact channel epochs into a single series: For each channel pair of interest, P, for condition C ∈ {word, non-word}, I utilized the following approach:

Let X_k, Y_k represent the time-series for epoch *k* for the two channels in P, respectively.

1) Normalize each epoch, $X_{k,t}$, by the transformation $X_{k,t} = X_{k,t} - \frac{\bar{X}}{\sigma}$ where $\bar{X}$ and σ are the sample mean and standard deviation, respectively, computed over all epochs.

2) Average the normalized epochs to obtain a single representative series: $X_t = \frac{1}{N}\sum_{k=1}^{N} X_{k,t}$

Y_t is computed similarly.

Time-domain averaging of multiple epochs retains frequency components that are phase-locked to the stimulus event (thus phase-locked across epochs) while reducing the effect of transient, spurious oscillations, presumably unrelated to the event. The latter should have random phase offsets between successive epochs and would thus be averaged out. Figure 2 shows the ECoG data for A1 and M1 channels after normalization and averaging.

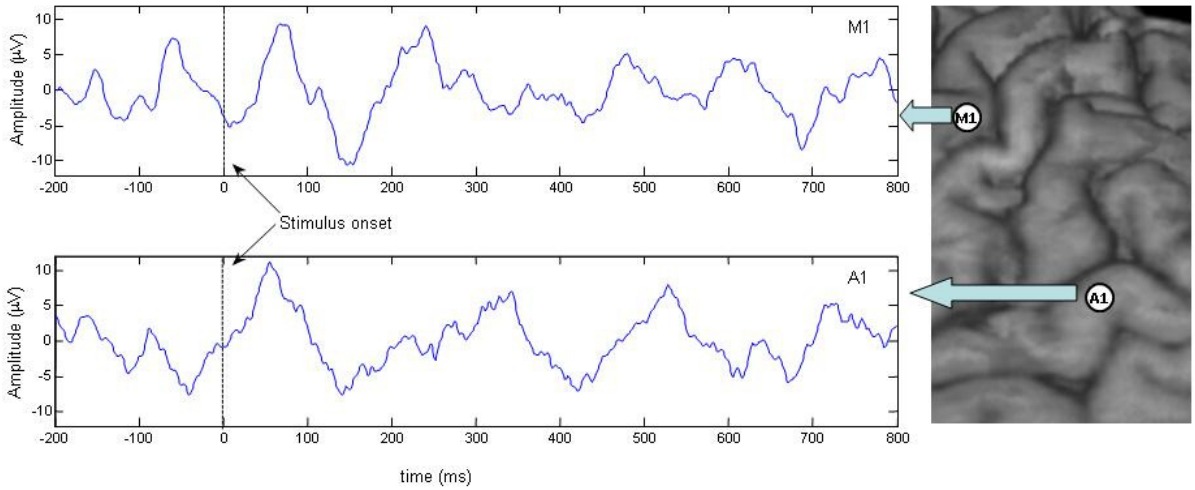

Figure 2. Normalized, averaged ECoG time-series for electrodes 58 (top) and 62 (bottom). Dashed line represents stimulus onset.

Observing the raw data plots, we note that the series does not appear to be stationary. While global trend does not appear to be present (this would have been removed by the 3 Hz high-pass filtering combined with the normalization procedure), there is significant local trend and a clear oscillatory effect is visible. Since I am primarily interested in analyzing linear interactions shortly after the stimulus, the data was truncated to the range 0 ms pre-stimulus to 400 ms post-stimulus for the remaining computations. The a.c.f. and c.c.f. estimates for A1 and M1 series are shown in Figure 3. Note that lags are in milliseconds rather than number of samples. Since $\rho_{XY}(k) = \rho_{YX}(-k)$, the plot for $\rho_{YX}(k)$ is not shown (it is simply the reflection about the y-axis of $\rho_{XY}(k)$). One can readily see that the a.c.f. is not "coming down" to zero quickly indicating that values at time *t* are highly correlated with multiple past values. This could be a result of local trend and/or cyclic/oscillatory behavior, both of which appear to be present in the series

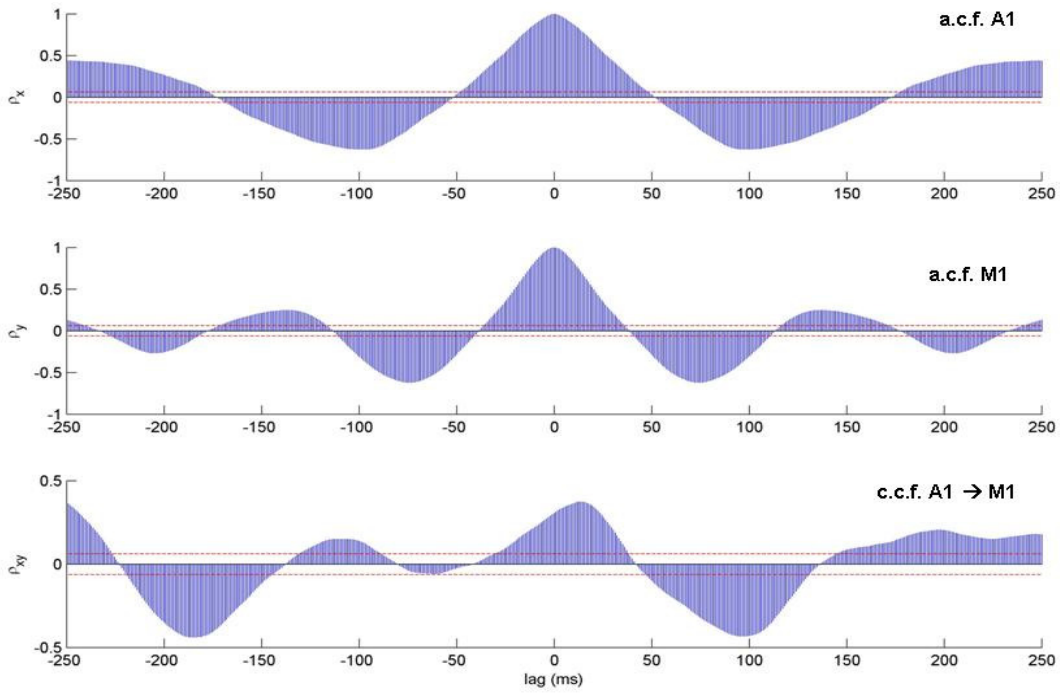

Figure 3. Estimates of autocorrelation for $\boldsymbol{X_t}$ (top), $\boldsymbol{Y_t}$ (mid), and cross-correlation E[$\boldsymbol{X_t}$,$\boldsymbol{Y_{t+k}}$] (bottom). Lags are in milliseconds. Dashed lines represent the ~95% confidence intervals obtained by ±2/√N where N is the length of series (Chatfield, 158).

(see Figure 2). The low-frequency cyclic effects in the autocorrelation function suggest a strong low-frequency periodic component in the data.

As an aside, it is worthwhile to note that that high autocorrelations in either series can spuriously inflate the variances of the cross-correlation estimates, making it difficult -- indeed dangerous -- to interpret the c.c.f. (Chatfield, 159). For example, observing the c.c.f. above, we might be inclined to hazard that the large c.c.f. peak at the 95ms lag indicates a directed 'propagation' effect at that latency from auditory to motor cortex. However, the prudent analyst would be reluctant to make such claims, noting that the a.c.f. estimates *also* have large peaks at that (and nearby neighboring) lags possibly inflating the cross-covariance estimates. The additional fact that there are several, large, wide peaks in the c.c.f. further frustrates interpretation.

An estimate of the power spectrum was computed using the method described in Brillinger, 1989, (and similar to periodogram smoothing procedures outlined in Chatfield, chapter 7) in which we take the

average of a set of periodogram estimates computed over *L* non-overlapping segments of the series. In order to get 1Hz resolution, I chose the number of fft points to be 1024. I chose a segment length of $4\sqrt{N}$ where N = length of the series. One can also obtain 95% confidence intervals on the spectrum using a procedure similar to that in Chatfield, ch7.5, and those are plotted in dotted lines bounding the spectrum. The spectrum for the non-differenced series is shown in Figure 4. The linear slope in the logarithmic scale (indicating an exponentially decreasing spectrum), along with the many peaks and troughs indicate that there is a great deal of overlapping oscillatory structure in the data, providing further evidence for it's non-stationarity. Some interesting peaks include those in the beta-range (13-30Hz).

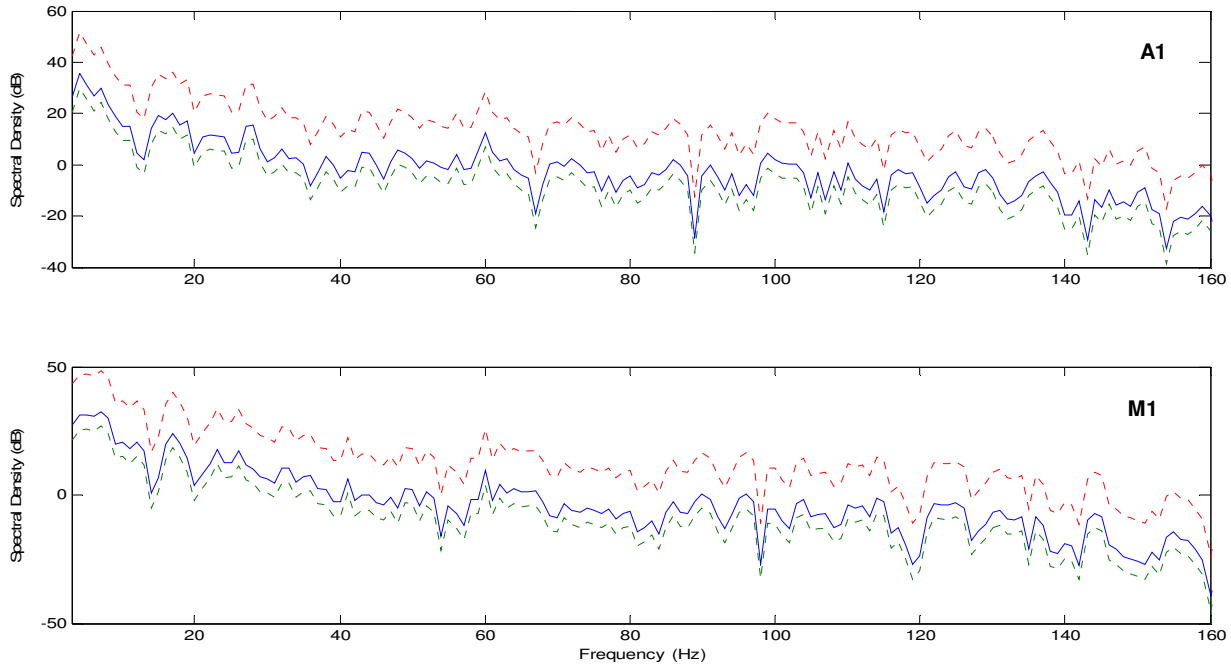

Figure 4. Log-linear Power spectrum (dB) for the non-differenced time series. 95% confidence intervals are dashed.

The above analyses all point to our series being non-stationary and thus it is appropriate to apply some differencing procedure to transform the data into a more stationary series prior to AR modeling.

3.3. Differencing

A commonly used transformation for improving the stationarity of a series is the first-difference operator, $\nabla X_t = X_t - X_{t-1}$. This operator can be re-applied iteratively *d* times. Seasonal differencing is

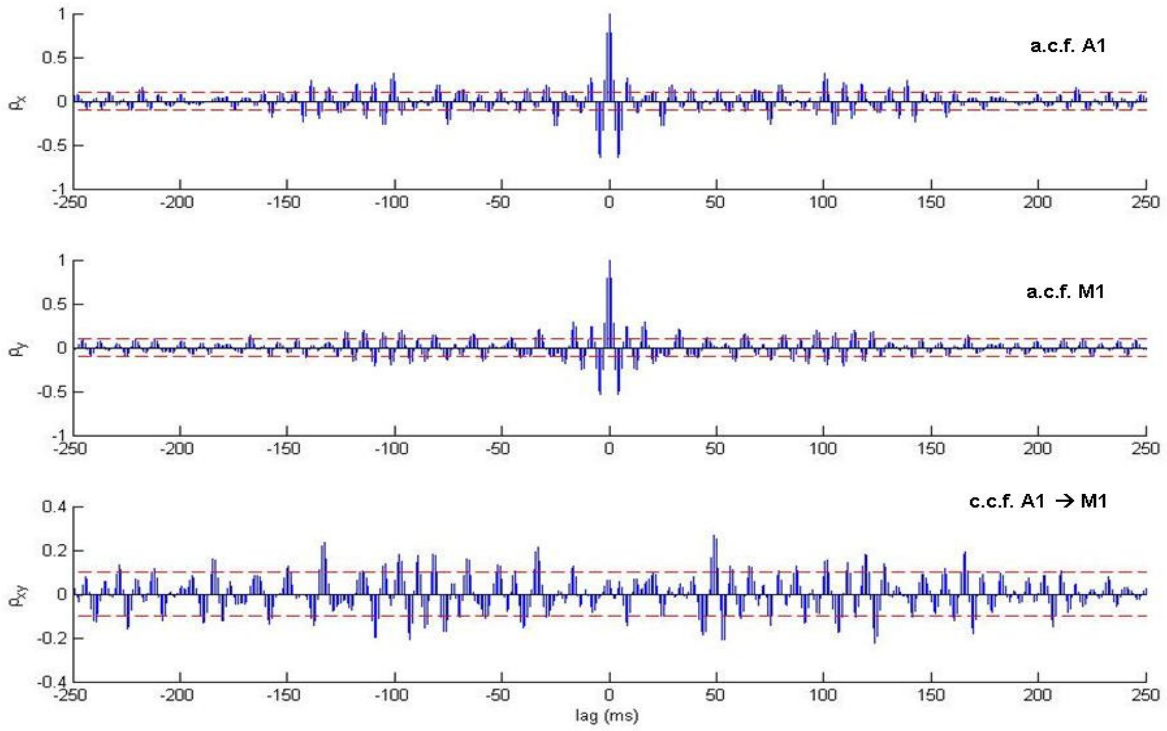

Figure 5. Autocorrelation and cross-correlation for second-order differenced series.

another useful operator, which simply takes the difference lag to be *s* rather than 1, where *s* is the period of a known periodic component of the data. However, our series are not seasonal and the periodicity seems to be due to many interacting periodic components rather than just one or two, thus a seasonal differencing operator may not be appropriate (I could not find a suitable seasonal lag for this series).

I chose to linearly detrend the series with a least-squares straight line fit, and then apply a second-order difference operator $\nabla^2 X_t = \nabla\nabla X_t$. The autocorrelations and cross-correlations are shown in Figure 5. Note that the a.c.f. now tails off much more quickly, but still contains some cyclic variation. QQ-plots of residuals versus normal distribution (Figure 6) show that the data is approximately normal.

3.4. Autoregressive Prewhitening

At this point, the stationarity condition seems sufficiently satisfied for application of autoregressive model fitting. An AR(49) model was fit to the input series using stepwise least squares estimation. Schwarz's Bayesian Criterion (SBC) (also known as Bayesian Information Criterion and similar to the

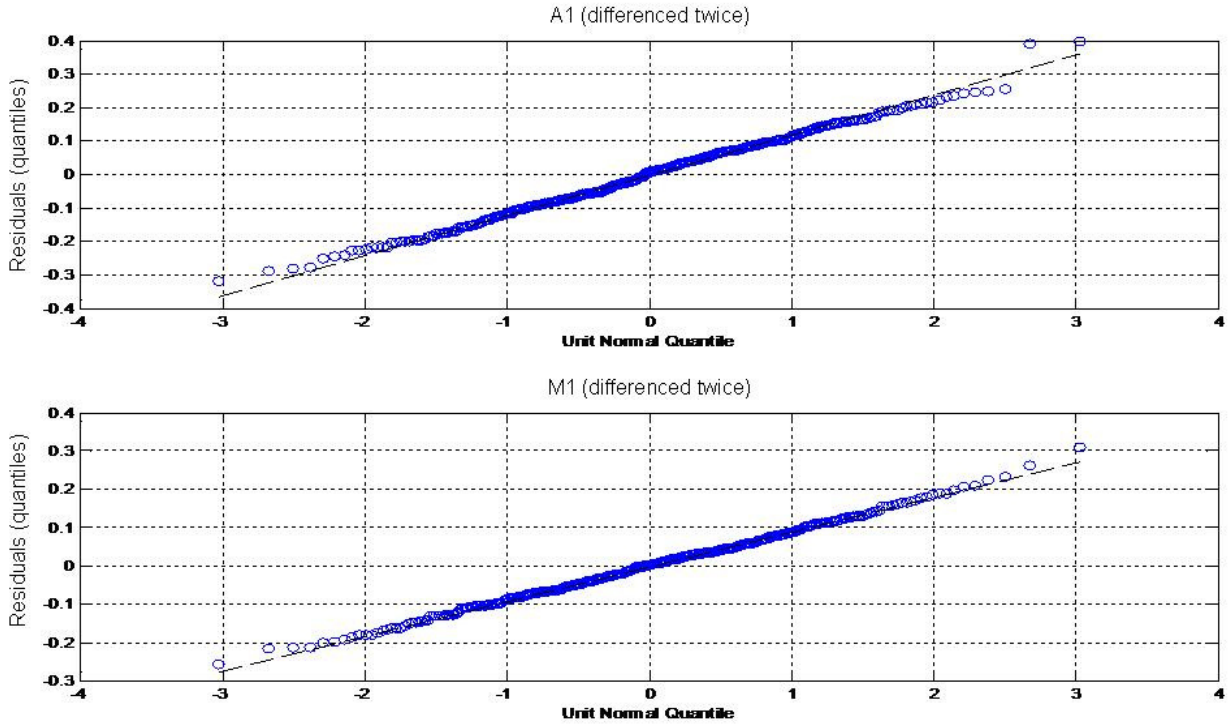

Figure 6. QQ-plot of residuals after differencing each series versus quantiles for normal distribution.

Aikaike Information Criterion) was used to automatically select the optimal model order. Model orders of this magnitude are not unusual for series containing a great deal of complex structure. The high sampling rate (1024 Hz) allows for a wider range of frequency information in the data which may necessitate the estimation of more parameters to obtain a suitable AR fit. The model fit was further evaluated by analyzing the a.c.f. and spectrum of the residual. The a.c.f. and c.c.f. are shown in Figure 7. If the residual was white noise we could expect approximately 5 out of 100 coefficients to lie outside the 95% confidence interval (CI). Our residuals show only 3 out of 100 coefficients outside the 95% CI, suggesting that the residuals are approximately white and the AR model is a good fit for the data. It's also interesting that the second series (M1) is also substantially whitened; indicating that the AR model also fits this series well. This might be due to structural similarity between the two series (which we might expect if the corresponding cortical regions are functionally related). Figure 8 shows the power spectra along with the cross spectrum computed on the residuals. Note that the spectrum of the input, X_t is *very* small (~ -60 dB) and rather flat suggesting that the data has indeed been reduced to white noise.

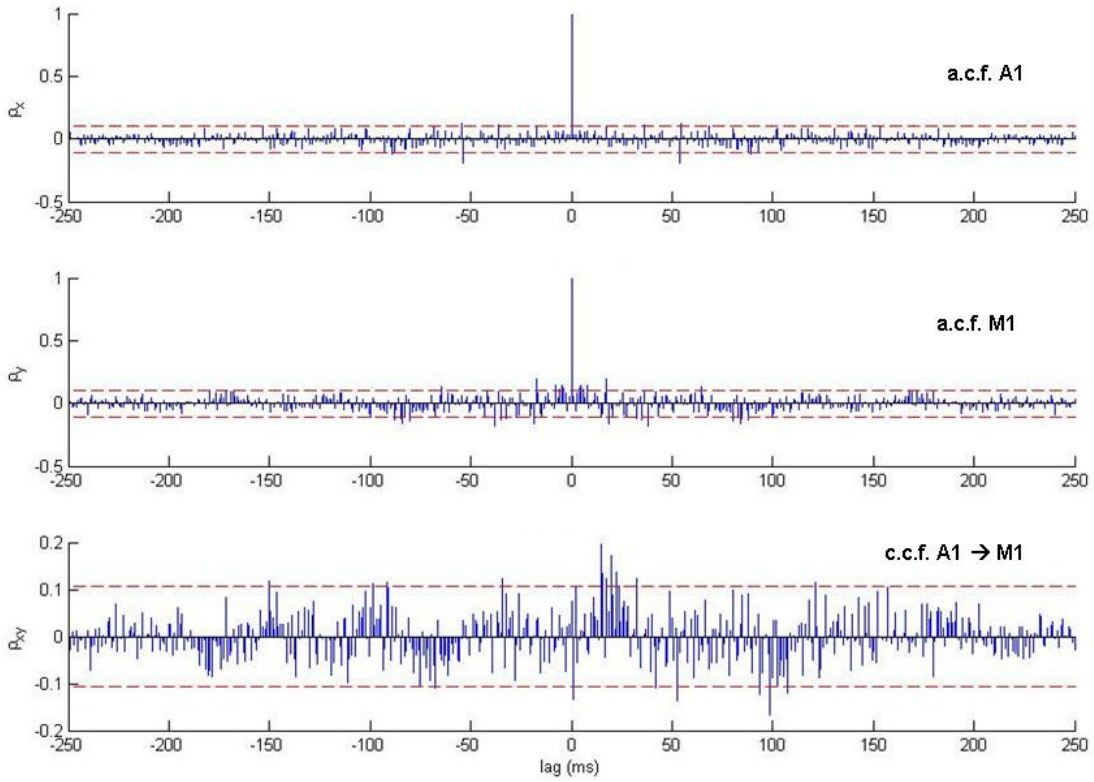

Figure 7. a.c.f. and c.c.f. for residuals of an AR(49) model fitted to the differenced data.

4. RESULTS

4.1. Cross-correlation

Now that we have properly whitened the data, we can look at the cross-correlation with critical eyes. Although there are a number of coefficients outside the 95% confidence interval, one might note immediately that there is a prominent peak at approximately 20ms. This cluster of coefficients spans the range 15-22 ms, well within the 10-36 ms propagation latencies estimated by direct cortico-cortical stimulation [4] and MEG peak-peak estimation [3]. Referring back to the cross-correlation between the spectra for the non-whitened, non-differenced data (Figure 3), one might note that there was a large peak around this range. However, it was difficult to ascertain whether this was a true directional effect or simply an artifact of high autocorrelation in the input series. Furthermore, the high values of neighboring coefficients along with the presence of other larger peaks made it difficult to determine if this was the primary delay of system.

4.2. Coherence, Phase, and Gain

Magnitude squared coherency (coherence), gain, and phase were all computed on the prewhitened data. These are shown in 9. From the gain diagram, it appears that this system is acting as a low-pass filter, in particular boosting the 15-30 Hz (Beta) frequency range. The coherence likewise shows a significant peak at 20Hz. The phase spectrum shows a negative slope in the low frequency range providing further evidence of a directed linear delay system with primary auditory cortex as the input and pre-motor cortex as the output. Fitting a linear least-squares approximation to the phase estimates at corresponding to significantly Beta-coherent frequencies yields a delay of approximately 10 ms from A1 to M1. While this is slightly less than the estimated cross-correlation latencies (15-20 ms), it is nevertheless within the propagation latency window of 10-36 ms estimated in [3, 4]. I further found that the relationships shown above did not hold if I computed the above estimates on *pre-stimulus* data (-200

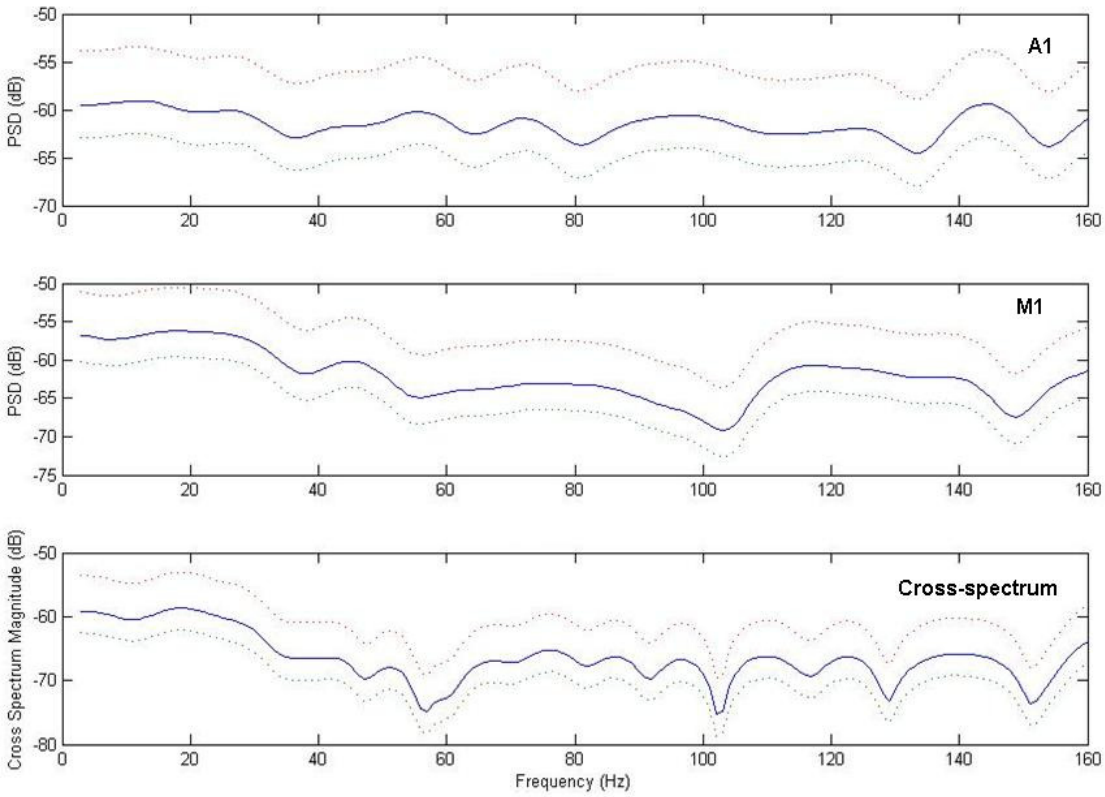

Figure 8. Power spectra $P_x(f)$ (top), $P_y(f)$ (mid) and co-spectrum (real part of cross-spectrum), $P_{xy}(f)$.

to 0 ms pre-stim). Those results are omitted for brevity.

5. DISCUSSION AND FUTURE WORK

In this paper I have carried out cross-correlation, coherence, gain, and phase estimates on time-domain averaged, autoregressively prewhitened ECoG data during a language comprehension task. The results suggest the presence of a functionally-specific linear delay system with primary auditory cortex projecting to primary motor cortex at an approximate delay of 10-20 ms. Furthermore, this interaction appears most prominent within the Beta frequency range, suggesting that motor-auditory network ensembles oscillating at ~20 Hz may be implicated in language processing.

The motor theory of language suggests that language comprehension takes place by activation and coupling of multiple motor subprograms. These programs may be activated by external stimuli such as a phonemic auditory pattern, and it seems reasonable to assume that auditory cortex may project

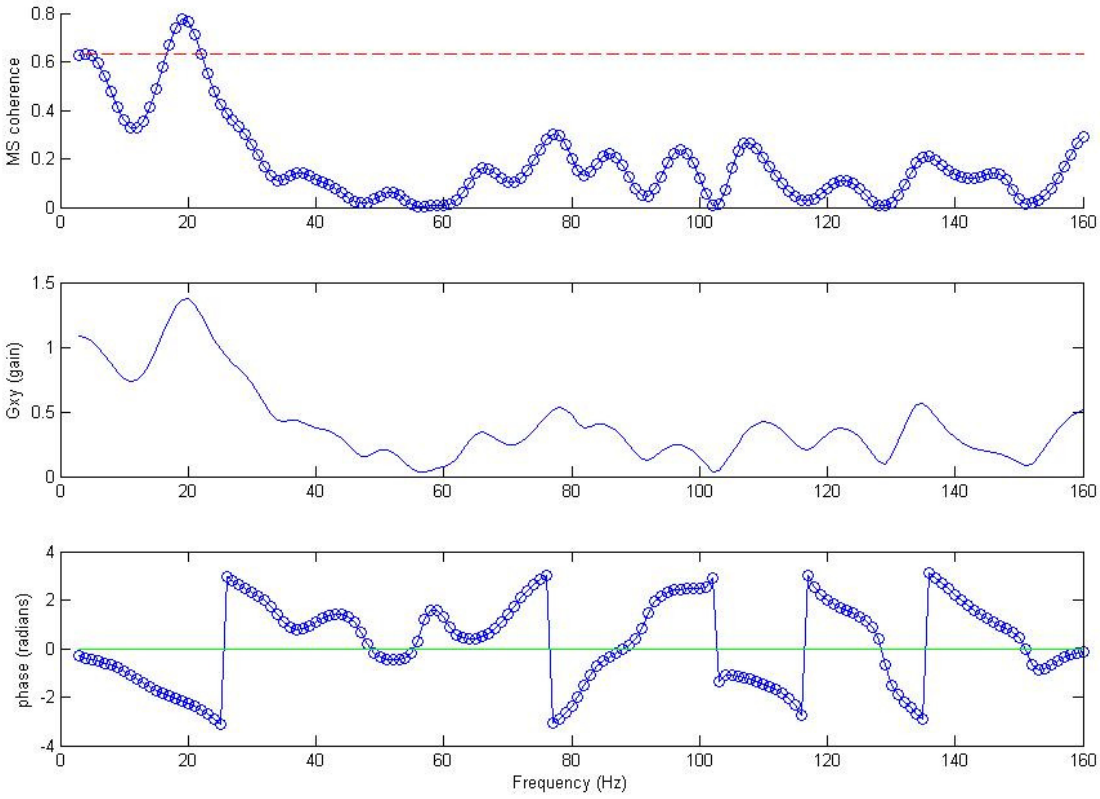

Figure 9. Magnitude squared coherency (top), Gain (mid), and Phase (bottom) across frequency. The dashed line in the coherence plot indicates ~ 95% confidence intervals computed as $1-(1-\alpha)^{1/(L-1)}$ where L is the number of periodograms used in estimating the coherence (Brillinger, 1989).

information directly to pre- and primary motor cortex. Beta oscillations have been implicated in large-scale sensorimotor network synchronization [5] and motor processing [6] and it is not unreasonable to posit that Beta oscillations may likewise be implicated in auditory-motor interaction during language comprehension. Naturally, these findings will need to be verified in multiple subjects. Additional work will include further exploration of parameter fitting for the linear delay model as well as the use of multivariate vector autoregressive (MVAR) models to account for feedback in the system.

6. REFERENCES

1. Fuhrmann, G., Mullen T., Suppiah, S., Soltani, M., Edwards E., Canolty, R., Sarang, D., Kirsch, H., Barbaro, N., Knight, R.T. (2007). *Functional Connectivity and Information Flow in the Human Brain during Language Processing: Evidence from ECoG data*. Poster. Cognitive Neuroscience Society Annual Conference, New York, N.Y.

2. Allott, Robin Michael (1994) *Motor theory of language origin: The diversity of languages*, in Wind, Jan and Jonker, Abraham and Allott, Robin and Rolfe, Leonard, Eds. *Studies in Language Origins Volume 3*, pages pp. 125-160. John Benjamins.

3. Pulvermüller, F. (2005). *Brain mechanisms linking language and action*. Nature Reviews Neuroscience 6, 576-582

4. Riki Matsumoto, Dileep R. Nair, Eric LaPresto, Imad Najm, William Bingaman, Hiroshi Shibasaki and Hans O. Lüders (2004) *Functional connectivity in the human language system: a cortico-cortical evoked potential study.* Oct;127(Pt 10):2316-30. Epub 2004 Jul 21.

5. Brovelli A, Ding M, Ledberg A, Chen Y, Nakamura R, Bressler SL (2004) *Beta oscillations in a large-scale sensorimotor cortical network: directional influences revealed by Granger causality*. Proc Natl Acad Sci USA 101:9849-9854. Epub 2004 Jun 9821.

6. Baker SN, Kilner JM, Pinches EM, Lemon RN (1999) *The role of synchrony and oscillations in the motor output*. Exp Brain Res 128:109-117.