Functional connections between auditory cortical fields in humans revealed by Granger causality analysis of intra-cranial evoked potentials to sounds: Comparison of two methods

Hiroyuki Oya [a,1], Paul W.F. Poon [a,b,*], John F. Brugge [a,2], Richard A. Reale [a,3], Hiroto Kawasaki [a,4], Igor O. Volkov [a,5], Matthew A. Howard III [a,6]

[a] *Department of Neurosurgery, University of Iowa College of Medicine, Iowa City, IA 52242, USA*
[b] *Department of Physiology, National Cheng Kung University, College of Medicine, Tainan, Taiwan*

Abstract

Knowledge of neural interactions amongst cortical sites is important for understanding higher brain function. We studied such interactions using Granger causality (GC) to analyze auditory event-related potentials (ERPs) recorded directly and simultaneously from two physiologically identified and functionally interconnected auditory areas of cerebral cortex in human neurosurgical patients. Two methods of GC analysis were used and the results compared. Both approaches involved adaptive autoregressive modeling but differed from each other in other ways. Results obtained by using the two methods also differed. Fewer false-positive results were obtained using the method that suppressed the ERP non-stationarity and that expressed the GC as the sum of model coefficients, which suggests that this is the more appropriate approach for analyzing ERPs recorded directly from the human cortex.
© 2006 Elsevier Ireland Ltd. All rights reserved.

* Corresponding author at: Department of Physiology, National Cheng Kung University, College of Medicine, Tainan, Taiwan. Tel.: +886 6 235 3535x5458; fax: +886 6 236 2780.

E-mail addresses: hiroyuki-oya@uiowa.edu (H. Oya), ppoon@mail.nkcu.edu.tw (P.W.F. Poon), john-brugge@uiowa.edu (J.F. Brugge), reale@physiology.wisc.edu (R.A. Reale), hiroto-kawasaki@uiowa.edu (H. Kawasaki), igor-volkov@uiowa.edu (I.O. Volkov), matthew-howard@uiowa.edu (M.A. Howard III).

[1] Tel.: +1 319 384 9846; fax: +1 319 353 8737.
[2] Tel.: +1 319 353 7082; fax: +1 319 353 8737.
[3] Tel: +1 608 263 5851; fax: +1 608 263 5929.
[4] Tel.: +1 319 384 8239; fax: +1 319 353 8737.
[5] Tel.: +1 319 335 8968; fax: +1 319 384 4830.
[6] Tel.: +1 319 356 8468; fax: +1 319 353 8737.

1. Introduction

Auditory cortex of the human is located on the superior temporal gyrus of the temporal lobe. It is made up of multiple fields that are thought to be interconnected and organized into three hierarchical processing levels referred to as the auditory 'core', 'belt' and 'parabelt'. This organizational model, which was originally derived from anatomical and physiological studies in non-human primates, is seen operating in a hierarchical fashion to process complex acoustic signals, such as communication sound and speech (reviewed by Kaas and Hackett, 2000; Rauschecker, 1998). Although this model has been posited for human cortex (Binder et al., 2000; Wessinger et al., 2001), there has been little direct experimental evidence for the cortico-cortical connec-

Fig. 1. Two views of human brain specimen showing the locations of the posteriolateral superior temporal field (PLST) on the superior temporal gyrus and the auditory core area on Heschl's gyrus (HG). (A) Lateral view of the surface of the right cerebral hemisphere showing PLST. (B) HG and surrounding cerebral cortex viewed from above after removing overlying parietal lobe. Arrows indicate the functional connectivity between HG and PLST.

tions required to support such a processing scheme. The anatomical tract-tracing methods that have been used so effectively in mapping auditory cortical fields and their connectivity in the living monkey brain cannot be used in humans. An alternative method of tracing auditory cortico-cortical pathways in the human brain involves focal electrical stimulation of one cortical site while systematically mapping the resultant evoked activity from distant sites (Brugge et al., 2003, 2005; Greenlee et al., 2004; Howard et al., 2000; Liegeois-Chauvel et al., 1991; Matsumoto et al., 2004). Using this approach we have shown functional connectivity between auditory cortex on mesial Heschl's gyrus (HG), which we consider to be core cortex, and an auditory area on the lateral surface of superior temporal gyrus, which we refer to as the posterior lateral superior temporal field (PLST) and interpret to be a part of the auditory belt or parabelt (Brugge et al., 2003, 2005; Howard et al., 2000). Fig. 1 shows on a human brain specimen the location of these two auditory fields, with arrows indicating the reciprocal functional connectivity between them.

Electrical mapping of functional connectivity patterns in temporal cortex provides important information

on timing, topography and direction of neural transmission, but leaves open the question of functional interactions presumably taking place between the interconnected fields under normal listening conditions. One approach to gaining this information is the so-called Granger causality (GC, Granger, 1969; Freiwald et al., 1999). Granger causality is defined in statistical terms and expressed as predictability: one stochastic process is causal to another if at a given time point the predictability of the second process is improved by including measurements from the immediate past of the first. When applied to interactions between distant brain sites, knowledge about the activity recorded from one cortical location has been used to predict activity recorded simultaneously at a second location (Brovelli et al., 2004; Chen et al., 2004; Ding et al., 2000; Hesse et al., 2003; Kaminski et al., 2001; Liang et al., 2000).

In the present study we measured during passive listening in humans the directional interactions that occurred between the core auditory cortex on posteromesial HG and auditory cortex on posterior lateral superior temporal gyrus, areas which have been shown previously to be functionally interconnected (Brugge et al., 2003, 2005; Howard et al., 2000). Data were auditory event-related potentials (ERPs) recorded directly from the cortex of neurosurgical patients. ERPs belong to the class of non-stationary signals (Oppenheim and Schafer, 1975) in which both mean voltage and variance tend to vary over time. Time-variant approaches to estimate GC are required to account for this non-stationary property. Two time-variant GC methods have been employed in the past to study functional interactions in the brain. One (Method 1) analyzed ERPs recorded from the scalp (Hesse et al., 2003) while the second (Method 2) analyzed local field potentials recorded intra-cranially (Ding et al., 2000; Kaminski et al., 2001). Both methods involve an adaptive autoregressive model but differ from each other in two important respects. First, Method 2 but not Method 1 pre-processes data to suppress signal non-stationarity. As we will demonstrate, skipping the pre-processing step can result in spurious causality responses. Second, Method 1 uses a ratio of error-variance produced by bi-variant modeling, whereas Method 2 employs the sum of coefficients on the off-diagonal coefficient matrices. Although both methods can be considered causality measures in the Granger sense, taking a ratio and interpreting its results needs to be done with care. At the start of our work it was unclear to us which approach would best apply to our near-field ERPs, hence we carried out GC analyses on the same dataset using both approaches and compared the results obtained.

2. Materials and methods

Subjects were five neurosurgical patients undergoing diagnosis and treatment of intractable epilepsy. All gave written informed consent and the protocols were approved by the University of Iowa Institutional Review Board. Depth electrodes were surgically inserted into HG while grid electrodes (8 × 8 array, 5 mm inter-electrode distance) were implanted directly on the pial surface of the superior temporal gyrus. Reference electrodes were either a surface contact placed on

inferior temporal gyrus or attached to the skull at the midline. Electrode recording sites were localized using intra-operative photographs and 3D MRI (for further methodologic details see Howard et al., 1996, 2000). Electrodes were placed on the right (non-dominant) hemisphere on four subjects and on the left (dominant) hemisphere of the fifth.

Synthesized consonant-vowel sounds (e.g. /ba/, duration ~250 ms), click trains (5 clicks, 10 ms inter-click interval) or 300 ms tone bursts (2 kHz, 5 ms rise/fall time) were used as stimuli. The sounds were delivered repetitively (2 s intervals)
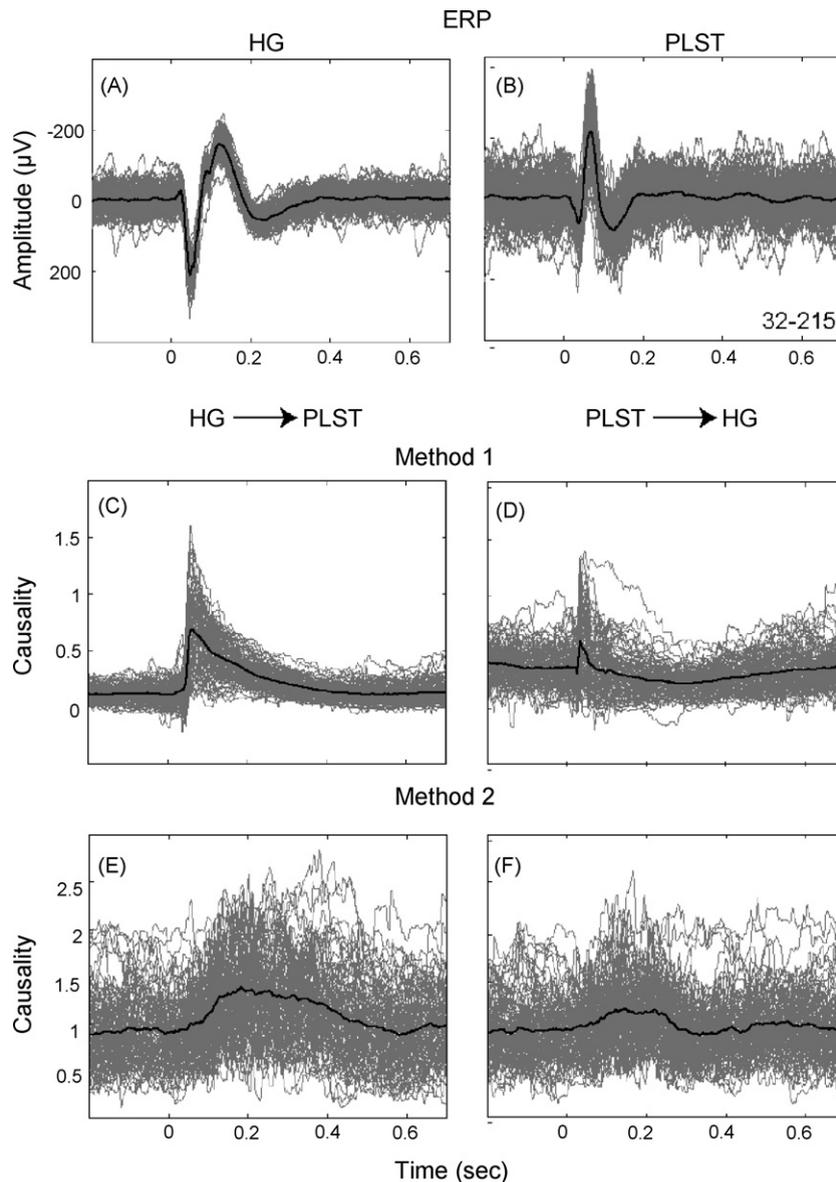
Fig. 2. Comparison of two methods of GC analysis for one subject. (A and B) Averaged ERPs (dark lines) and single trial records (gray lines) obtained from simultaneous intracranial recording in posteromesial HG and PLST in response to a brief click train (*n* = 100 trials). (C–F) Averaged Granger causality (GC) profiles (dark lines) and single trial data (gray lines) obtained from ERP results shown above by Method 1 (from Hesse et al., 2003) and Method 2 (from Ding et al., 2000; Kaminski et al., 2001). In this and other figures, the stimulus onset is at time zero. Stimulus duration is approximately 40 ms for click trains. Direction of causal influence is given above the respective GC analyses.

at a comfortable suprathreshold level in blocks of 100 identical trials through binaurally placed insert-earphones. During recording sessions patients were usually awake and resting in bed, and were instructed to listen passively to the sounds.

Signals recorded simultaneously from surface and depth electrode contacts were amplified, filtered (1–1000 Hz), digitized (2 kHz) and stored for offline analysis. Line interference was reduced by a multi-tapering method (Mitra and Pesaran, 1999). To speed up computation, electrical potentials were first re-sampled to 250 Hz. Typically one site in posteromesial HG and one on posterolateral superior temporal gyrus exhibits an ERP having an overall amplitude larger than ERPs around it. In each of the five subjects studied we subjected the ERPs recorded from these two sites to GC analyses.

In Method 2, but not in Method 1, a pre-processing step was introduced prior to autoregressive modeling in order to compensate for the non-stationary nature of the ERP signal (Ding et al., 2000). In this step, the average ERP waveform was first subtracted from the individual recordings. The resultant waveforms were then divided, point-by-point, by the ensemble variance. The result was that the ensemble mean and variance became constant over time.

The first step in the application of our autoregressive model was to choose how much of the past history of the ERP recorded at one channel should be used in predicting the influence on the ERP recorded from the second channel. This time window is referred to as the 'order' for the ERP signals. To choose the most appropriate order a Bayesian information criterion (BIC) function (Schwarz, 1978) was first calculated for all 100 recording trials in a dataset over the full length (900 ms) of each trial. For each trial the time window associated with the minimum in the BIC function was found. The distribution of time windows for each 100-trial dataset was then plotted, and the time window corresponding to the 90th percentile of that distribution was taken as the optimal order for that dataset. This approach tended to avoid under-estimation while taking into account outliers. A median order value of 3 (inter-quartile range 2) was obtained (eight experiments), which corresponds to 12 ms. This order is only slightly less than the one used by Ding et al. (2000) in his studies of local field potentials in monkey.

The second step was the introduction of a Kalman filter, which was used to track the time-variant state of the system adaptively in a recursive way. When the disturbances and initial state are Gaussian the Kalman filter is optimal, and it remains an optimal 'linear' estimator even when the Gaussian assumption is dropped (Appendix A). We attempted to improve the tracking capability of the Kalman filter by manipulating the magnitude in the change of the autoregressive coefficients using a normalized innovation-squared criterion (Appendix A.1). The estimated model coefficients and disturbance covariance allowed estimation of the inter-relationship between recording channels. This inter-relationship between ERPs was then expressed either as the ratio of the disturbance error variance with or without an ERP from the second recording site (Method 1, Appendix B.1), or as the sum of coefficients on the off-diagonal coefficient matrices (Method 2, Appendix B.2). The

causality measure derived from Method 2 is also referred to as 'direct causality' (Kaminski et al., 2001). Repeating this procedure for all time points in a response trial produced a GC time waveform (causality profile). For each trial a moving time-window (corresponding to the order of the autoregressive model) was adopted for calculating the model coefficients over the peri-stimulus period of 900 ms (200 ms before and 700 ms after stimulus onset). The results of averaging 100 trials were compared. Further details describing the two approaches used are found in Appendices A and B.

3. Results

The polyphasic ERP recorded from posteromesial HG could be distinguished from that recorded from PLST on the posterolateral superior temporal gyrus (Howard et al., 2000). As seen in Fig. 2 A and B, the average ERP from each site exhibits early deflections associated with incoming afferent activity and later, and larger, deflections which an be attributed to intra-cortical interactions. Considerable time-dependent trial-to-trial
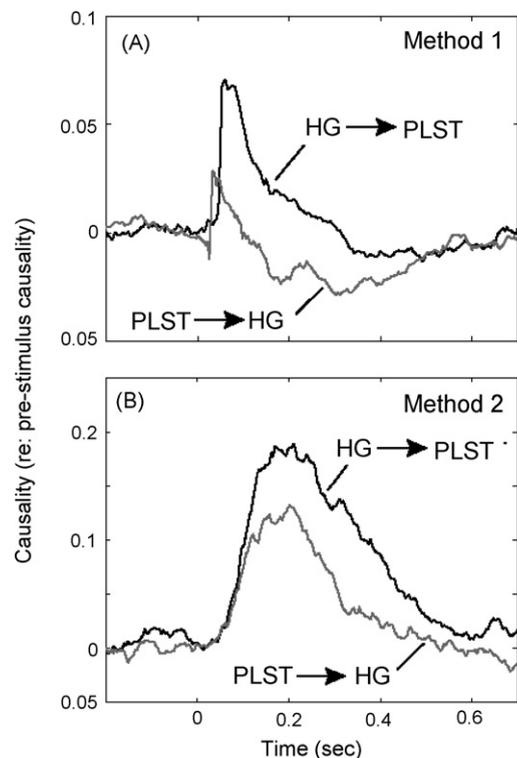
Fig. 3. Comparison of two methods of GC analysis averaged for eight experiments in five subjects. Stimuli were click trains (two experiments), tone bursts (two experiments) or a consonant-vowel sounds (four experiments). Directions of causal influence shown are from HG to PLST (black line) and from PLST to HG (gray line). The average pre-stimulus causality has been subtracted to better reveal the stimulus-related GC changes.

variation is also seen, and in this case is demonstrably larger in PLST as compared to posteromesial HG. The GC profiles computed from these waveforms using Methods 1 and 2 are shown in Fig. 2C–F. Those on the left represent influences in the direction from HG to PLST, whereas those on the right show influences in the opposite direction. Fig. 3 compares the averaged causality profiles from eight experiments in our five subjects derived from the two methods. The averaged profiles closely resemble those shown for the one subject in Fig. 2 and further reveal a discrepancy between the two

approaches we compared. In both Methods 1 and 2, GC influences computed in the direction from HG to PLST and from PLST to HG appeared as incremental changes time-locked to the onset of the stimulus and in close temporal registration with the ERP. Both methods yielded GC profiles that showed the duration of the influence to be direction dependent. Causality time profiles differed, however, between the two methods. Method 1 produced a sharp transient peak in causality for both directions, which decreased compared to the prestimulus condition. Method 2, on the other hand, yielded a gradual and more
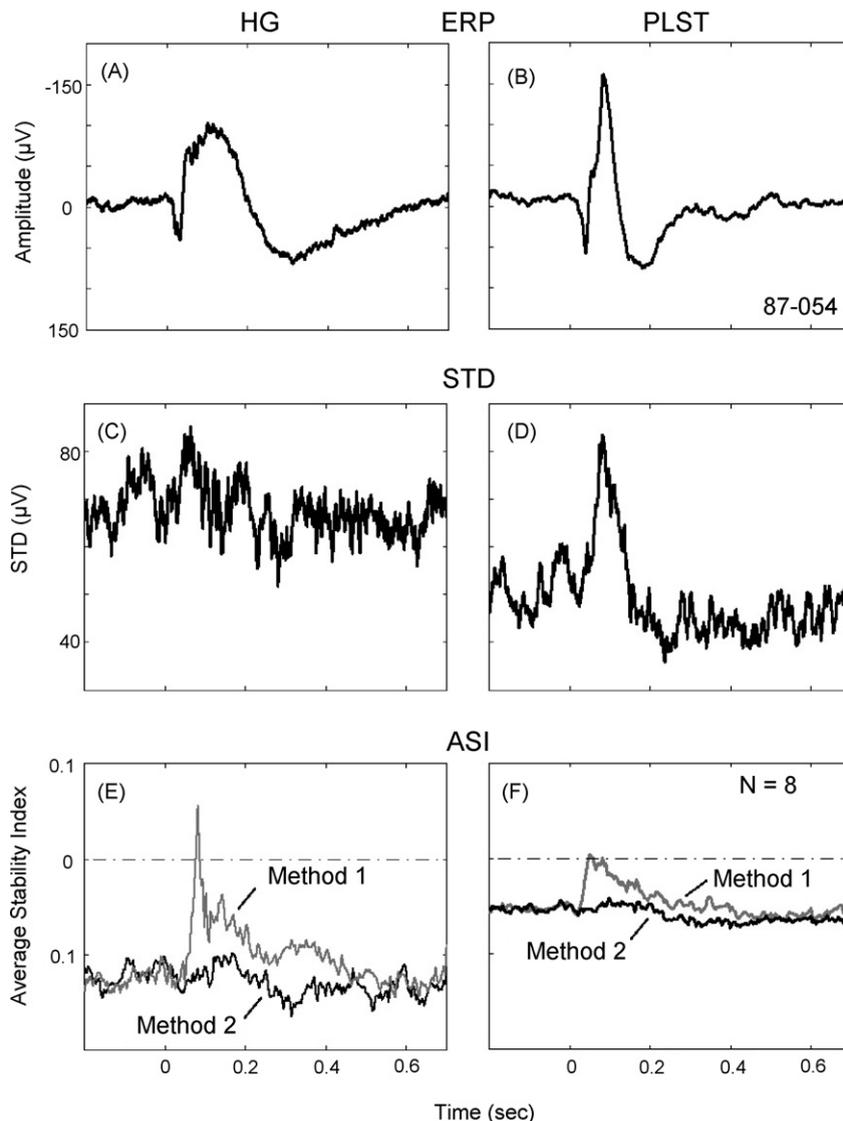
Fig. 4. Effect of data pre-processing on the GC. (A and B) Averaged ERPs obtained from simultaneous intracranial recordings in posteromesial HG and PLST in response to a brief click train (*n* = 100 trials). (C and D) Time course of ensemble standard deviation (S.D.) of the recording presented in A and B, respectively. S.D.s are not constant over the time course of ERP and could vary across recording sites. (E) Averaged stability index (ASI) of the same dataset shown in A and C with pre-processing (black line; Method 2) or without pre-processing (gray line; Method 1). (F) Same as E showing the average results of eight datasets from five subjects.

sustained increase in causality for both directions and never dropped below zero (by definition).

To study further the possible cause(s) of the discrepancy between the two methods employed, we examined more closely the time-dependent variation of ERPs from another subject shown in Fig. 4A and B. Fig. 4C and D shows the accompanying time-dependent variance. The
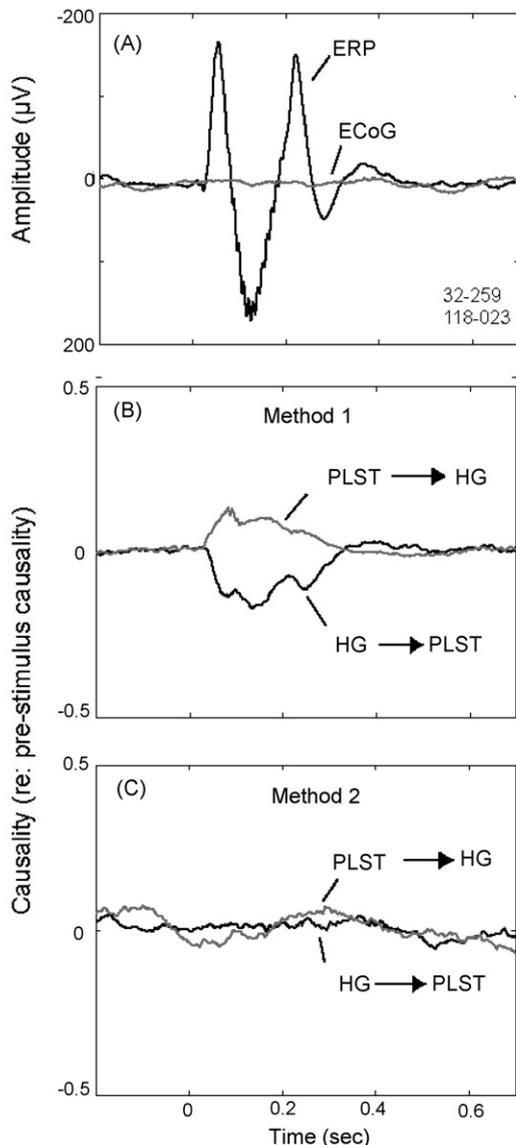
Fig. 5. Negative control. (A) Average ERPs from HG in one subject (dark line) in response to a consonant-vowel sound (/da/) superimposed on an average electrocorticographic (ECoG) signal from PLST (grey line) of a different subject (*n* = 100 trials). (B and C) Averaged GC profiles computed by both methods from data shown in A (*n* = 100 trials). The average pre-stimulus causality has been subtracted to better reveal the stimulus-related GC changes. Black lines indicate causal influence from HG to PLST and gray lines indicate PLST to HG. Spurious results seen using Method 1 are not seen using Method 2.

addition of the data pre-processing step in the analysis strongly suppressed the non-stationarity as revealed by the 'average stability index' (ASI) shown in Fig. 4E and F. The ASI is the logarithm of the largest (absolute) eigenvalue of the companion matrix of the autoregressive coefficients averaged across trials. To the extent the ASI remains below zero the autoregressive model is considered stationary. As applied to the data in Fig. 4A and B, the gray line in Fig. 4E shows ASI obtained by Method 1 whereas the black line show the ASI obtained by Method 2. Method 2 resulted in suppression of non-stationarity that was almost complete as reflected in a nearly constant and below-zero ASI over the analysis period of 900 ms. The similar picture emerged when the ASI, obtained by the two methods, was averaged over all datasets of the five subjects (Fig. 4F).

Although the data presented suggest that spurious results of GC such as those shown in Figs. 2 and 3 could arise from non-stationarity in the ERP signal, other factors could be involved here as well. We further subjected the two analysis methods to a pair of signals in which the causality was necessarily zero: one of the pair was an ERP recorded on HG to a speech sound and the other an ongoing electrocorticographic (ECoG) signal recorded in silence from PLST of a different subject (Fig. 5A). Method 1 showed noticeable stimulus-locked changes in GC (Fig. 5B) while Method 2 yielded GC profiles that remained relatively constant over the recording period (Fig. 5C).

4. Discussion

Granger causality was used to analyze functional influences between two auditory cortical fields known to have functional connections. We compared two methods of computing GC based on ERPs recorded directly from the cortex of human neurosurgical patients. Method 1 was applied originally by Hesse et al. (2003) to ERPs recorded from the human scalp. Method 2 was first described by Ding and his colleagues (Ding et al., 2000; Kaminski et al., 2001) in studies of ERPs recorded intracranially. We found that the results of Method 2 more faithfully represented our ERP data than did the results of Method 1.

The discrepant results we obtained by comparing the two approaches could be partly due to the non-stationary nature of the ERP signal. Hesse et al. (2003) did not test for non-stationarity in their GC analysis of scalp-recorded ERPs. On the other hand, their ERPs had signal-to-noise ratios that were smaller than those of the intracranial ERPs studied by Ding et al. (2000) and Kaminski et al. (2001) and by us. Hence non-stationarity

may have had a far smaller impact on the outcome of their GC analysis. Expressing GC as a ratio (Method 1) instead of a sum of coefficients (Method 2) could also create greater sensitivity to signal non-stationarity and hence lead to spurious results. Estimating the noise covariance is not always easy as the driving noise level may change over time. Thus taking the ratio of noise covariance can have a direct impact on the results obtained.

It seems clear that three steps were crucial in the GC analysis of our ERP data. First, the Kalman filter was used to track the autoregressive coefficients more efficiently. Second, the sum of model coefficients was used to express GC, which proved to be more stable than the ratio of disturbance of noise variance. Third, and most importantly, data were pre-processed to remove the deterministic evoked response, which reduced signal non-stationarity.

Our preliminary results using Method 2 from a limited number of subjects suggest that auditory core cortex on HG and auditory area PLST on the posterolateral superior temporal gyrus interact in a reciprocal way during passive listening. In addition, the data indicate that the time course of the interactions is direction-dependent. These findings are consistent with results of electrical stimulation tract-tracing experiments in human temporal cortex (Brugge et al., 2003, 2005; Howard et al., 2000).

Acknowledgements

Appendix A

We have organized the appendix into two parts: (A) background of autoregressive modeling common to the two methods, and (B) technical details of each method.

A.1. Time variant multi-dimensional autoregressive modeling

A common approach for modeling uni-variate time series is the autoregressive (AR) model, which is simply a linear regression of the current value of the time series against one or more prior values (p) of the series. The value of p is called the order of the AR model. The general equation for a stationary first-order linear-predictive time-varying autoregressive model is described by following equation: $y(t) = \xi y(t-1) + \varepsilon(t)$ where $y(t)$ is the observed time series, ξ is the autoregressive coefficient and $\varepsilon(t)$ is a white noise sequence, the so called driving noise. If $y(t)$ is a multivariate vector and coefficients ξ are time dependent, the model becomes a time-variant vector autoregressive model (VAR) and has been applied to investigate correlation of electrical activity between different brain sites (Arnold et al., 1998; Möller et al., 2001). The time-variant VAR we employed in this study was represented in state-space form, which can be estimated by Kalman filter recursion as described below.

A.1.1. State-space formulation of the autoregressive model

The m-dimensional vector autoregressive model of order p can be expressed in following state-space form:

$$X(t) = F(t)\ X(t-1) + W(t)$$

$$Y(t) = H(t)\ X(t) + E(t)$$

where $F(t)$ is the transition matrix with dimension ($m^2 p \times m^2 p$), and $H(t)$ is the measurement matrix with dimension ($m \times m^2 p$), defined as

$$H_t = I_m \otimes [Y_{t-1}, Y_{t-2}, \ldots, Y_{t-p}]$$

I_m is the identity matrix of dimension m and $\otimes$ represents Kronecker's matrix product. $W(t)$ is the ($m^2 p \times m^2 p$)-dimensional zero-mean white process-noise sequence with the following covariance matrix.

$$E[W_t W_t^{\mathrm{T}}] = Q(t)$$

$E(t)$ is the ($m \times m$)-dimensional zero-mean white observation-noise sequence with the following covariance matrix:

$$E[E_t\ E_t^{\mathrm{T}}] = R(t)$$

X_t are state variables. In this application, X_t are the autoregressive coefficient matrices at time t and are defined as follows:

$$X_t = \mathrm{Vech}[A_1(t), A_2(t), \ldots, A_p(t)]$$

where $A_n(t)$ is the n-lag autoregressive coefficient matrix at time t, Vech is the vectorization operator, and Y_t is the m-dimensional data vector (observation vector) at time t. The coefficient matrix X_t that represents the underlying structure of the system is hidden and what we can observe is the time series $Y(t)$. The aim here is to estimate X_t from the observed $Y(t)$. Here, we assume the coefficients evolve according to a random walk. So $F(t)$ is the identity matrix.

A.1.2. Recursive adaptive estimation of time-varying state by means of a Kalman filter

A Kalman filter is commonly used for tracking time-variant states (e.g., navigation, radar target tracking and time series prediction in economics) (Ljung, 1987). Kalman filter recursion can be written as follows:

$$K_t = \widehat{\Sigma}_{t|t-1} H(t)^{\mathrm{T}} [H(t)\widehat{\Sigma}_{t|t-1} H(t) + R(t)]^{-1}$$

$$\widehat{\Sigma}_{t|t} = \widehat{\Sigma}_{t|t-1} - K_t H(t)\widehat{\Sigma}_{t|t-1}$$

$$\hat{X}_{t+1|t} = F(t)[\hat{X}_{t|t-1} + K_t(Y(t) - H(t)\hat{X}_{t|t-1})]$$

$$\widehat{\Sigma}_{t+1|t} = F(t)\widehat{\Sigma}_{t|t} F(t) + Q(t)$$

$$\hat{X}_{0|-1} = \hat{X}_0$$

$$\widehat{\Sigma}_{0|-1} = \Sigma_0$$

where K_t is the Kalman gain matrix, $\widehat{\Sigma}_{t|t}$ the *a priori* covariance matrix of the estimation error and $\widehat{\Sigma}_{t+1|t}$ the *a posteriori* covariance matrix of the estimation error. To set the observation noise covariance $R(t)$, we estimated with the following procedure.

State-noise covariance $Q(t)$ determines the speed of tracking. We gave equal tracking capability to all coefficients, i.e. $Q(t) = \sigma I$, where I is the identity matrix with dimension ($m^2 p$). In general, σ can be time-variant and adaptively changed. For doing this, we monitored the smoothed normalized innovation squared (NIS) function. NIS can be thought of as the estimated variance of standardized prediction errors:

$$\mathrm{NIS}(t) = \mathrm{NIS}(t-1) - a[\mathrm{NIS}(t-1) - v(t)^{\mathrm{T}} S(t)^{-1} v(t)]$$

where $S(t)$ is the covariance of innovation produced by the Kalman recursion:

$$S(t) = H(t)\widehat{\Sigma}_{t|t-1} H(t) + R(t)$$

Under the hypothesis that the filter is consistent, NIS has a chi-square distribution with L degrees of freedom (L is the effective window length, i.e., $L = m/\alpha$. Here we set α at 0.03). We started σ at $10^{-3.5}$, and changed σ according to the following criterion. If NIS(t) fell within 90–95% confidence intervals, σ for the next time point became $10^{-2.5}$; if NIS(t) exceeded 95%, σ became $10^{-2.0}$. The initial level of σ was set to $10^{-3.5}$.

$\hat{X}_{t|t} = \hat{X}_{t|t-1} + K_t(Y(t) - H(t)\hat{X}_{t|t-1})$ is the optimal estimator of $X(t)$ based on the observations up to $Y(t)$. The noise term not explained by the model is calculated as

$$E(t) = Y(t) - H(t)\hat{X}_{t|t}$$

Its covariance $V(t)$ is estimated by following the exponentially discounted average:

$$V(t) = V(t-1) - \gamma[V(t-1) - E(t)^{\mathrm{T}} \widehat{\Sigma}_{t|t} E(t)]$$

Here, γ is the discounting factor, which was set to 0.03. This $V(t)$ was used for $R(t)$ in the Kalman recursion shown above.

This produces estimators of the autoregressive model coefficients and the innovation covariance matrix for each time point. For this estimation procedure, no stationary restriction was imposed on the autoregressive coefficient matrix. The stationarity of the estimated autoregressive model can be checked by examining, point by point, the characteristic roots of the time-variant coefficient matrix. In both Methods 1 and 2, we performed autoregressive modeling on each ERP trial, and the average time profiles of GC obtained for different datasets were compared.

Appendix B. Inferring directed linear dependencies: technical details

B.1. Method 1

In a vector autoregressive model, linear dependence and feedback are used to estimate Granger causality (GC, Geweke, 1982). For the following bi-variate autoregressive model:

$$Y(t) = \sum_{i=1}^{p} \Gamma_i(t) Y(t-i) + \varepsilon(t) \text{ with } \mathrm{var}[\varepsilon(t)] = \Omega(t)$$

and the uni-variate autoregressive model, in which one of the time series is omitted from the system:

$$y_1(t) = \sum_{i=1}^{p} \Gamma_i^1(t) y_1(t-i) + u_1(t)$$

$$y_2(t) = \sum_{i=1}^{p} \Gamma_i^2(t) y_2(t-i) + u_2(t) \quad \text{with}$$

$$\mathrm{var}[u_1(t)] = \Theta_1(t) \quad \mathrm{var}[u_2(t)] = \Theta_2(t)$$

Here the matrices Γ are the autoregressive coefficient matrices consisting of $x_{t|t}$ and the time-variant variance of the disturbance terms in the system; p is the model order, ε and u are estimated as follows:

$$\Omega(t) = \Omega(t-1) - a[\Omega(t-1) - \varepsilon(t)\varepsilon(t)^{\mathrm{T}}],$$

$$\Theta_1(t) = \Theta_1(t-1) - a[\Theta_1(t-1) - u_1(t)u_1(t)^{\mathrm{T}}],$$

$$\Theta_2(t) = \Theta_2(t-1) - a[\Theta_2(t-1) - u_2(t)u_2(t)^{\mathrm{T}}]$$

Here, we set the smoothing constant (a) to 0.03 in order to suppress noisy variations in the results. Time-variant GCs in the opposite direction between channel-1 and channel-2 are, respectively, $GC_{1-2}(t)$ and $GC_{2-1}(t)$. These are defined as

$$GC_{1-2}(t) = \log |\Theta_2(t)|/\log |\Omega^{2,2}(t)|,$$

$$GC_{2-1}(t) = \log |\Theta_1(t)|/\log |\Omega^{1,1}(t)|$$

For the calculation of the causality, we fit the uni- and bi-variate models, as described in section A, with Θ and Ω replaced by $V(t)$ in the Kalman filter.

B.2. Method 2

First, the ERP individual sweeps were pre-processed according to Ding et al. (2000). Namely, the ensemble average waveform was subtracted from individual recordings followed by dividing the resulted waveforms by the ensemble variance point by point, so that the ensemble mean and variance became constant over time. Then, we applied bi-variate autoregressive modeling and the estimated coefficient matrices were used to infer the directed linear dependencies as described below.

In the autoregressive model in the previous section:

$$Y(t) = \sum_{i=1}^{p} \Gamma_i(t)Y(t-i) + \varepsilon(t)$$

The elements in $\mathbf{\Gamma}_i(\mathbf{t})$ represent linear time-lagged dependence (i: time lag). At a given time, lag i, the diagonal elements in $\mathbf{\Gamma}_i(\mathbf{t})$ represent self-connectivity and the off-diagonal elements inter-channel connectivity. For example, in the following first order autoregressive model with dimension of 2:

$$\begin{bmatrix} \text{sig1}(t) \\ \text{sig2}(t) \end{bmatrix} = \begin{bmatrix} a1 & a2 \\ a3 & a4 \end{bmatrix} \begin{bmatrix} \text{sig1}(t-1) \\ \text{sig2}(t-1) \end{bmatrix} + \begin{bmatrix} \varepsilon1(t) \\ \varepsilon2(t) \end{bmatrix}$$

$a1$ and $a4$ represent self-connectivity at lag 1, and $a2$ and $a3$ represent inter-channel connectivity at lag 1.

In the above coefficient matrix, the influence from sig1 at $t-1$ is expressed explicitly by the coefficients $a1$ and $a3$. So the ratio abs(a3)/(abs(a1) + abs(a3)) represents the relative strength of influence from sig1 to sig2 relative to the total output of sig1 at time $t-1$.

Normalized direct causality (N-DC) (Kaminski et al., 2001) is defined as follows:

$$N - DC(t)[\text{chA} \rightarrow \text{chB}] = \frac{\sum_{i=1}^{p} \text{abs}(\Gamma_{(B,A)}(i,t))}{\sum_{i=1}^{p} \text{abs}(\Gamma_{(:,A)}(i,t))}$$

where abs($\Gamma_{(:,\text{chA})}(i,t)$) represents the sum of elements (in absolute value) of Ath column in the coefficient matrix at time t for lag i, and abs($\sum_{i=1}^{p}\Gamma_{(B,A)}(i,t)$) represents the sum of Ath column (in absolute value), Bth row element in the coefficient matrix at time t for lag i. In the Granger sense, a non-zero causality means the existence of non-zero coefficients in the off-diagonal elements. The sign of the coefficient (positive or negative) has no direct implication on Granger causality. Furthermore, taking absolute values of the coefficients avoids the problem of cancellation which might have occurred in Method 1. Also, the non-normalized direct causality (DC) is defined as follows:

$$DC(t)[\text{chA} \rightarrow \text{chB}] = \sum_{i=1}^{p} \text{abs}(\Gamma_{(B,A)}(i,t))$$

In this method, we used the non-normalized DC as an index of causality measure, as it minimizes the effect of the denominator term for easier interpretation of results.

References

Arnold, M., Miltner, X.H.R., Witte, H., Bauer, R., Braun, C., 1998. Adaptive AR modeling of nonstationary time series by means of Kalman filtering. IEEE Trans. Biomed. Eng. 45, 553–562.

Binder, J.R., Frost, J.A., Hammeke, T.A., Bellgowan, P.S., Springer, J.A., Kaufman, J.N., Possing, E.T., 2000. Human temporal lobe activation by speech and non-speech sounds. Cereb. Cortex 10, 512–528.

Brovelli, A., Ding, M., Ledberg, A., Chen, Y., Nakamura, R., Bressler, S.L., 2004. Beta oscillations in a large-scale sensorimotor cortical network: directional influences revealed by Granger causality. Proc. Natl. Acad. Sci. USA 101, 9849–9854.

Brugge, J.F., Volkov, I.O., Garell, P.C., Reale, R.A., Howard, M.A., 2003. Functional connections between auditory cortex on Heschl's gyrus and on the lateral superior temporal gyrus in humans. J. Neurophysiol. 90, 3750–3763.

Brugge, J.F., Volkov, I.O., Reale, R.A., Garell, P.C., Kawasaki, H., Oya, H., Li, Q., Howard, M.A., 2005. The posteriolateral superior temporal auditory field in humans. Functional organization and connectivity. In: Konig, R., Heil, P., Budinger, E., Scheich, H. (Eds.), The Auditory Cortex—Toward a Synthesis of Human and Animal Research. Erlbaum, Mahwah, NJ, pp. 145–162.

Chen, Y., Rangarajan, G., Feng, J., Ding, M., 2004. Analyzing multiple nonlinear time series with extended Granger causality. Phys. Lett. A 324, 26–35.

Ding, M., Bressler, S.L., Yang, W., Liang, H., 2000. Short-window spectral analysis of cortical event-related potentials by adaptive multivariate autoregressive modeling: data preprocessing, model validation, and variability assessment. Biol. Cybern. 83, 35–45.

Freiwald, W.A., Valdes, P., Bosch, J., Biscay, R., Jimenez, J.C., Rodriguez, L.M., Rodriguez, V., Kreiter, A.K., Singer, W., 1999. Testing non-linearity and directedness of interactions between neural groups in the macaque inferotemporal cortex. J. Neurosci. Methods 94, 105–119.

Geweke, J., 1982. Measurement of linear dependence and feedback between multiple time series. J. Am. Stat. Assoc. 77, 304–313.

Granger, C.W., 1969. Investigating causal relations by econometric models and cross spectral models. Econometrica 37, 424–438.

Greenlee, J.D., Oya, H., Kawasaki, H., Volkov, I.O., Kaufman, O.P., Kovach, C., Howard, M.A., Brugge, J.F., 2004. A functional connection between inferior frontal gyrus and orofacial motor cortex in human. J. Neurophysiol. 92, 1153–1164.

Hesse, W., Möller, E., Arnold, M., Schack, B., 2003. The use of time-variant EEG Granger causality for inspecting directed interdependencies of neural assemblies. J. Neurosci. Methods 124, 27–44.

Howard, M.A., Volkov, I.O., Abbas, P.J., Damasio, H., Ollendieck, M.C., Granner, M.A., 1996. A chronic microelectrode investigation of the tonotopic organization of human auditory cortex. Brain Res. 724, 260–264.

Howard, M.A., Volkov, I.O., Mirsky, R., Garell, P.C., Noh, M.D., Granner, M., Damasio, H., Steinschneider, M., Reale, R.A., Hind, J.E., Brugge, J.F., 2000. Auditory cortex on the posterior superior temporal gyrus of human cerebral cortex. J. Comp. Neurol. 416, 76–92.

Kaas, J.H., Hackett, T.A., 2000. Subdivisions of auditory cortex and processing streams in primates. Proc. Natl. Acad. Sci. USA 97, 11793–11799.

Kaminski, M., Ding, M., Truccolo, W.A., Bressler, S.L., 2001. Evaluating causal relations in neural systems: Granger causality, directed transfer function and statistical assessment of significance. Biol. Cybern. 85, 145–157.

Liang, H., Ding, M., Nakamura, R., Bressler, S.L., 2000. Causal influences in primate cerebral cortex during visual pattern discrimination. Neuroreport 11, 2875–2880.

Liegeois-Chauvel, C., Musolino, A., Chauvel, P., 1991. Localization of the primary auditory area in man. Brain 114, 139–151.

Ljung, L., 1987. System Identification: Theory for the User. Prentice-Hall, Englewood Cliffs, NJ.

Matsumoto, R., Nair, D.R., LaPresto, E., Najm, I., Bingaman, W., Shibasaki, H., Luders, H.O., 2004. Functional connectivity in the human language system: a cortico-cortical evoked potential study. Brain 127, 2316–2330.

Mitra, P.P., Pesaran, B., 1999. Analysis of dynamic imaging data. Biophys. J. 76, 691–708.

Möller, E., Schack, B., Arnold, M., Witte, H., 2001. Instantaneous multivariate EEG coherence analysis by means of adaptive high-dimensional autoregressive models. J. Neurosci. Methods 105, 143–158.

Oppenheim, A.V., Schafer, R.W., 1975. Digital Signal Processing. Prentice-Hall, Englewood Cliffs, NJ.

Rauschecker, J.P., 1998. Cortical processing of complex sounds. Curr. Opin. Neurol. 8, 516–521.

Schwarz, F., 1978. Estimating the dimension of a model. Ann. Stat. 6, 461–464.

Wessinger, C.M., VanMeter, J., Tian, B., Van Lare, J., Pekar, J., Rauschecker, J.P., 2001. Hierarchical organization of the human auditory cortex revealed by functional magnetic resonance imaging. J. Cogn. Neurosci. 13, 1–7.