Supporting Information

Schoffelen et al. 10.1073/pnas.1703155114

SI Materials and Methods

Experimental Procedure and MEG Data Acquisition. A total of 102 native Dutch speakers (51 males), with an age range of 18–33 y (mean of 22 y), participated in the experiment. All participants were right-handed, had normal or corrected-to-normal vision, and reported no history of neurological, developmental, or language deficits. The study was approved by the local ethics committee (CMO, the local "Committee on Research Involving Human Participants" in the Arnhem–Nijmegen region) and followed the guidelines of the Helsinki declaration. Participants received monetary compensation for their participation.

The participants were seated comfortably in a magnetically shielded room and were instructed to read sequences of words (total number of 240, with 9–15 words per sequence), which were presented sequentially on a back-projection screen, placed in front of them. All words were presented at the center of the screen within a visual angle of 4°, in a black monospaced font, on a gray background using Presentation software (Version 16.0; Neurobehavioral Systems, Inc.). The vertical refresh rate of the LCD projector was 60 Hz. The sequences of words formed either well-formed sentences or consisted of a scrambled version of a sentence, where the word order was randomly shuffled. For the remainder we refer to these latter stimulus sequences as word lists. See Lam et al.(23) for more details about the stimulus material used. Sentences and word lists were presented in small blocks, of five sentences (or word lists) each, to a total of 120 stimuli per condition. To check for task compliance, in a random 10% of the word sequences they were followed by a yes/no question about the content of the previous sentence/word list.

MEG data were collected with a 275 axial gradiometer system (CTF). The signals were analog low-pass-filtered at 300 Hz and digitized at a sampling frequency of 1,200 Hz. The participant's head was registered to the MEG-sensor array using three coils attached to the participant's head (nasion and left and right ear canals). Throughout the measurement the head position was continuously monitored using custom software (26). During breaks the participant was allowed to reposition to the original position if needed. Participants were able to maintain a head position within 5 mm of their original position. Three bipolar Ag/AgCl electrode pairs were used to measure the horizontal and vertical electrooculogram and the electrocardiogram.

Artifact Rejection and Subtraction of Single-Trial Activity. All analyses were done with custom-written MATLAB scripts and FieldTrip (27). Data were initially epoched from −100 to 600 ms relative to word onset. Segments contaminated by artifacts due to eye movements, muscular activity, and superconducting quantum interference device jumps were discarded before further analysis. Next, we subtracted the event-related response from the single-trial data with the ASEO algorithm (28). The acronym ASEO stands for analysis of single-trial ERP and ongoing activity, and the aim of the application of this algorithm in this context was to attenuate the effects of evoked transients on the estimation (and subsequent interpretation) of GC (29). Transients in the signals violate the underlying assumption of stationarity and moreover may result in nonzero GC estimates, due to systematic latency differences of the peak of the transient signals across regions. Although such latency differences may reflect an actual interaction (where temporal precedence of a transient signal peak in region A, compared with region B, may be an indication that A is causing B), their spurious effect on the estimated GC is unwanted, if the aim is to interpret the frequency domain GC in terms of directed synchronized interactions. In our experimental setup, transient brain responses could not be avoided (as opposed to, for instance, ref. 7), and we developed a procedure (performed at the sensor level) to attenuate the effect of transient evoked components, combining the ASEO algorithm with a blind source separation technique (DSS) (30). In short, the ASEO algorithm models single-trial signals as a combination of ongoing activity and event-related components, where the latter are modeled as a set of "canonical" components, each with a trial-specific latency and amplitude. In a typical application (31), the single-trial estimates of latencies and amplitude are used as dependent variables for subsequent analysis. Here, however, we subtracted the single-trial evoked responses that were reconstructed from the latency and amplitude estimates, which results in a better account of ongoing activity, compared with the subtraction of a fixed average event-related response from each signal. DSS was used to iteratively unmix the sensor-level data into a set of components, where the DSS framework allows for the unmixing algorithm to capitalize on specific features of the requested components. Specifically, we applied an iterative procedure, where each iteration consisted of the following steps:

i) Estimation of the dominant DSS component using quasiperiodic averaging, which essentially extracts components with strong evoked transients, time-locked to word onset. This step yields a spatial map of mixing weights, describing for each MEG sensor the extent to which this component is present in the MEG signals, as well as an observations-by-time matrix of the component time series. Mathematically, the sensor-level data are modeled as:

$$[X_1\ X_2\ \ldots X_R] = a * [y_1\ y_2 \ldots y_R] + N,$$

where matrix X_r has channels in the rows and time points in the columns for trial r. Column vector a represents the mixing weights, and row vector y_r represents the single-trial time courses of the estimated component. For the implementation we used the software provided by the authors (30) (available from www.cis.hut.fi/projects/dss/), which is part of FieldTrip and can be deployed using `ft_componentanalysis`, with `cfg.method = 'dss'`

ii) Application of the ASEO algorithm to the single-trial time series of the estimated DSS component (vector y in our notational convention), yielding single-trial estimates of the evoked transients. The data model underlying the ASEO algorithm models each single-trial response as a combination of ongoing activity and a weighted combination of latency shifted event-related canonical signal components, as per the following equation:

$$y_r(t) = \sum_{n=1}^{N} \beta_m s_n(t - \tau_m) + z_r(t) = \tilde{y}_r(t) + z(t),$$

where $y_r(t)$ is the signal for trial r as a function of time t. $z_r(t)$ is the ongoing activity, and $s_n(t)$ is a canonical signal component, which is shifted by latency parameter τ and weighted by amplitude parameter β. The ASEO algorithm parameterizes the single-trial signals using an iterative approach in the frequency domain. For the implementation used in the current work we adapted the code that was kindly provided to us by Mingzhou Ding, University of Florida, Gainesville, FL, and which was used for the original implementation (28).

iii) Backprojection of the component's single-trial evoked transients $\tilde{y}_r$ to the MEG sensor level, using the spatial map of mixing weights, obtained in step *i*:

$$[\tilde{X}_1\ \tilde{X}_2\ \ldots\tilde{X}_R] = a * [\tilde{y}_1\ \tilde{y}_2 \ldots \tilde{y}_R].$$

iv) Subtraction of the back-projected evoked transients from the MEG sensor data, yielding MEG-sensor data that served as input data for the next iteration.

We performed five iterations (i.e., we removed five time-locked components). Removal of additional components did not affect the global field power appreciably (Fig. S3*E*). The different steps and the effect of this cleaning procedure are illustrated in Fig. S3.

Source Reconstruction and Parcellation of Source-Reconstructed Activity. After the cleaning of the sensor data with the combined DSS-ASEO procedure we performed source reconstruction using an LCMV (32). For this, we computed the covariance matrix between all MEG-sensor pairs, as the average covariance matrix across the cleaned single trial covariance estimates. This covariance matrix was used in combination with the forward model, defined on a set of 8,196 locations on the participant-specific reconstruction of the cortical sheet to generate a set of spatial filters, one filter per dipole location. Individual cortical sheets were generated with the Freesurfer package (version 5.1) (surfer.nmr.mgh.harvard.edu), coregistered to a template with a surface-based coregistration approach, using Caret software (brainvis.wustl.edu/wiki/index.php/Caret:Operations/Freesurfer_to_fs_LR, brainvis.wustl.edu/wiki/index.php/Caret:Download), and subsequently downsampled to 8,196 nodes, using the MNE software (martinos.org/mne/stable/index.html). The forward model was computed using FieldTrip's "singleshell" method (41), where the required brain/skull boundary was obtained from the subject-specific T1-weighted anatomical images.

Next, we applied an atlas-based parcellation scheme to further reduce the dimensionality of the data. To this end, we used the Conte69 atlas (brainvis.wustl.edu/wiki/index.php//Caret:Atlases/Conte69_Atlas), which provides a parcellation of the neocortical surface, based on Brodmann's cytoarchitectonic atlas, consisting of 41 labeled parcels per hemisphere. This parcellation scheme was further refined, breaking up the larger parcels into a set of subparcels, respecting the original boundaries (e.g., breaking up the middle temporal gyrus in smaller parcels along the anterior/posterior axis). This resulted in a parcellation scheme consisting of 191 parcels per hemisphere.

For each parcel, we obtained a parcel-specific spatial filter as follows. We concatenated the spatial filters of the vertices comprising the parcel, obtained a set of time courses of the event-related field at each parcel, and performed a principal component analysis on the result. We selected for each parcel the first two spatial components explaining most of the variance in the signal. We opted for this method, rather than averaging, because we used rank-reduced forward solutions (excluding the most noise sensitive dipole orientations), which might result in signal cancellation effects upon averaging, due to sign ambiguity of the resulting cardinal dipole orientations. For the parcels used the two dominant spatial components explained on average 90% of the signal variance within each parcel (range: 74–96%).

Preselection of the Connections Between Language-Relevant Areas. For the connectivity analysis we constrained ourselves a priori to a subset of connections between parcel pairs, using known "long-range" macroanatomical fiber pathways between parcels comprised of core language regions and the visual system as described in the literature (1, 9–11). This preselection was motivated by the fact that direct functional connections should be supported by direct anatomical connections. In addition, we allowed a priori for direct connections between neighboring nodes, which is a fair assumption given the characteristics of cortico-cortical connections observed in anatomical tracing studies (e.g., refs. 24 and 33), where local connections are abundant. We included intrahemispheric connections from both hemispheres and also included interhemispheric connections between homologous areas. The nodes were defined based on the labeling scheme of Brodmann, where each of these nodes could consist of one or more subparcels (where subparcels were defined as described above). In addition, nodes in temporal cortex were classified according to their position along the anterior–posterior axis (distinguishing anterior, middle, and posterior parts) and along the superior–inferior axis (distinguishing superior, middle, and inferior parts). Fig. 2*A* in the main text shows how the individual nodes were labeled. As major long-range fiber pathways we included the arcuate fasciculus (AF), the superior longitudinal fasciculus (SLF), the extreme capsule (EC), the uncinate fasciculus (UC), and the inferior fronto-occipital fasciculus (IFOF). The AF provides widespread connections between the temporal cortex (predominantly the middle and superior temporal gyri) and various frontal areas (BA44/45/6/9). The SLF connects frontal areas (notably BA44) with posterior superior temporal and parietal areas. The EC connects frontal areas with the middle part of superior and middle temporal gyrus. The UC connects frontal areas with the temporal pole, and the IFOF connects frontal areas with occipital areas. Connections between directly adjacent parcels were excluded for further analysis to reduce spurious estimates of connectivity due to spatial leakage of source reconstructed activity. The selection scheme resulted in 4,350 connections between pairs of parcels, which notably consisted of a sparse subset of all possible pairwise connections between the 156 parcels used for the GC analysis.

GC Computation and Statistical Evaluation of Overall Network Topology. For computational efficiency we computed the spectral representation of the signals at the sensor level and projected this into source space, using the parcel-specific spatial filters. The spectral representation of the signals was obtained using the fast Fourier transform in combination with multitapers (using 5-Hz smoothing) on the time domain data from 200 until 600 ms after word onset. The sensor-level Fourier-transformed data were projected into source space, and for each pair of parcels we computed the cross-spectral density matrix. Subsequently we performed nonparametric spectral matrix factorization for each pair of parcels, followed by computation of GC (34, 35).

We used a blockwise approach (35) in combination with nonparametric spectral factorization (34). Rather than using a parametric approach, which requires the estimation of multivariate autoregressive models, and subsequently uses the Fourier transforms of the model coefficients (along with variance estimates of the residuals) to obtain an estimate of frequency-resolved GC, we used a nonparametric technique, which is implemented in FieldTrip and has successfully been used before (6, 42).

The factorization algorithm decomposes a (by definition conjugate symmetric) cross-spectral density matrix, into a (frequency-specific) asymmetric spectral transfer matrix, and a frequency independent symmetric matrix, according to

$$S(\omega) = H(\omega)\Sigma H^*(\omega),$$

where $S(\omega)$ is the cross-spectral density as a function of frequency ω, $H(\omega)$ is the spectral transfer matrix, and Σ is a square symmetric matrix, representing the equivalent of the residuals' covariance matrix in a parametric estimation context. In a bivariate context, that is, where each of the signals consists of a univariate time series, and thus the cross-spectral density S is a 2×2 matrix for each frequency bin, GC from signal 1 to signal 2 is defined as

$$GC_{1\to 2}(\omega) = \ln\left(\frac{S_{22}(\omega)}{S_{22}(\omega) - \left(\Sigma_{11} - \frac{\Sigma_{21}^2}{\Sigma_{22}}\right) abs(H_{21}(\omega))^2}\right),$$

where the subscripts in the variables reflect the row and column indices for the individual elements of the 2×2 matrices. In a blockwise context, where each of the signals is a multivariate (in our case, a bivariate) time series, GC from signal 1 to signal 2 is defined, according to ref. 35, as

$$GC_{1\to 2}(\omega) = \ln\left(\frac{|S_{22}(\omega)|}{|\tilde{H}_{22}(\omega)\Sigma_{11}\tilde{H}_{22}(\omega)|}\right),$$

where the subscripts in the variables reflect the row and column indices of the individual blocks of the 4×4 matrices, |X| denotes the determinant of a matrix, and $\tilde{H}$ is a modified version of the spectral transfer matrix, defined as

$$\tilde{H}(\omega) = H(\omega)P^{-1},$$

where P is defined as:

$$\begin{bmatrix} 1 & 0 & -\Sigma_{12}\Sigma_{22}^{-1} \\ 0 & 1 & \\ 0 & 0 & 1 \quad 0 \\ 0 & 0 & 0 \quad 1 \end{bmatrix}.$$

For the implementation we used custom-written code (to improve computational efficiency), which was adopted from the FieldTrip functions `sfactorization_wilson`, `sfactorization_wilson2×2`, and `ft_connectivity_granger`.

In addition to computing GC, we also computed GC based on the source-projected Fourier transform of time-reversed data, where time reversal is essentially equivalent to complex conjugation of the Fourier coefficients, to distinguish "weak" asymmetries from "strong" asymmetries, as described by Haufe and coworkers (13) and Haufe et al. (36). Essentially, a weak asymmetry is an apparent directional interaction between a pair of network nodes, which is the consequence of a difference in SNR across nodes (14) and difficult to avoid when the signals consist of a linear mixture of underlying sources (37). We compared GC with reverse GC and selected only parcel pairs for subsequent analysis for which the parametric null hypothesis of the means (across subjects) could be rejected at a P value < 0.05, corrected for multiple comparisons (one-sided t test, with Bonferroni correction). This reduced the number of connections that were used for subsequent analysis from 4,350 to 713. Next, we evaluated the topology of this resulting network by quantifying the node degree for each of the 156 parcels involved, identifying "hubs" for inflow and outflow (Fig. 1). We quantified the probability of observing the computed node degree under the null hypothesis of the 713 connections' being a random subset of the originally included 4,350 connections, using a permutation test. Using Bonferroni correction, a P value of $0.05/(2*156) = 1.6 \times 10^{-4}$ was considered significant (each of the 156 parcels was tested twice, once for the degree for inflow and once for the degree for outflow).

As an important control analysis, we computed, across parcels, the Spearman's rank correlation between the inflow and outflow degree on the one hand and signal variance, the norm of the spatial filter, and the SNR on the other hand. The norm of the spatial filter corresponds with an estimate of the projected noise, and the ratio between the signal variance and the spatial filter's norm corresponds with an SNR estimate. The motivation for this analysis is to check whether there is a relationship between the node degree and simple univariate signal(-to-noise) properties, which may give rise to spurious inferences about the directionality of estimated interactions (14). Specifically, assuming the worst, one could hypothesize that parcels with a large degree of inflow (outflow) also show on average a low (high) signal(-to-noise), when comparing across parcels. The results of this control analysis are shown in Table S1. Based on this analysis, which did not reveal any significant correlations, we argue that the observed patterns of node degree in the brain network for language are not consequences of systematic differences in univariate signal properties.

NMF and Network Visualization. We explored the network topology by performing NMF with sparsity constraints, using an open-source toolbox (38). Specifically, we used the `sparsenmfnnls` routine from this toolbox, to group-concatenated data matrix X according to

$$X \approx AY,$$

where X, A, and Y are nonnegative matrices and A and Y aim to minimize the following quantity:

$$X - AY_F^2 + \eta A_F^2 + \lambda \sum_{i=1}^{N} y_{i1}^2,$$

where η and λ are sparsity parameters.

Because typically the outcome of NMF depends on the random starting conditions, we repeated the factorization multiple times ($n = 40$), and used a hierarchical clustering procedure, as implemented in the Icasso toolbox (40) (research.ics.aalto.fi/ica/icasso). In particular, we used the `hcluster` function, using as a similarity matrix a matrix that was defined as

$$1 - corr\left(\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_{40} \end{bmatrix}\right),$$

where $corr(Y)$ is the sample correlation matrix between all pairs of rows in Y, and where the matrix Y was constructed as a concatenation of the network topologies extracted by each of the 40 runs of the NMF algorithm. We fixed the number of clusters to 20. From each of the clusters, we extracted the component that on average had the highest correlation with the other members of that cluster, to obtain the spatial topology and spectral content of that cluster.

The purpose of this analysis is to describe the reconstructed connectivity data as a low-dimensional mixture of network components, each of which with a subject-specific spectral profile. This technique has successfully been applied to sensor-level MEG-signals before (39). We opted for sparse NMF, because the nonnegativity constraint facilitates the interpretation of the components, as opposed to, for example, a statistical independence constraint as applied in independent component analysis. This is because GC is strictly nonnegative. The data matrix that was subjected to the factorization algorithm was constructed by concatenating across subjects GC spectra, normalized for the SD per subject. The columns in this matrix reflect the individual connections (across subjects and frequencies), and the rows in this matrix (number of frequency bins times number of subjects) reflected the connections for a given frequency bin and subject. Typically, the outcome of NMF is dependent on the number of components (which has to be chosen a priori) and of the initial random starting conditions. We explored a range of "number of components" but settled on the number 20, because this number provided a reasonable balance between providing a small number of easily interpretable components, while at the same time maintaining a good separation between subnetworks.

The outcome of this procedure consisted of two matrices. One matrix represents the network component spatial fingerprints, quantifying for each of the edges its relative contribution to the network components. The other matrix contains for each network component a subject-specific spectral profile, quantifying the subjectwise relative and frequency-specific contribution to the network components. For visualization purposes we assigned each of the edges to a unique network component based on its relative weight. Subsequently, the different aspects of the network components were depicted as follows. To obtain spatial maps of the nodes (i.e., anatomical parcels) participating in a particular component, we summed across each node's contributing edges the outflow and inflow separately and displayed these onto an inflated representation of the cortical sheet, using hcp-workbench (www.humanconnectome.org/software/connectome-workbench.html).

To visualize the connections between the parcels, for each pair of parcels we averaged the connection weights across all pairs of subparcels that constituted the parcel pair. Connections were drawn as directed arrows, where the thickness of the lines reflects the overall weight of the connection. The spectral profiles of the network components were visualized as the median across subjects and their IQR. For display purposes we ordered the components according to their dominant regions for outflow. Fig. 2 in the main text shows the components that are made up predominantly by connections between language-relevant cortical parcels.

Components that are made up predominantly by connections between visual cortical parcels as well as components with spatially very diffuse connections are displayed in Fig. S1.

As a sanity check, we compared the NMF-based results with K-means clustering scheme, which in theory should give very similar results (43, 44). The results of this comparison are displayed in Fig. S4. We used the MATLAB built-in `kmeans.m` function, with `cityblock` as distance metric. Overall, both clustering approaches yield very similar patterns, where the K-means clusters were topologically less sparse than their corresponding NMF-based counterparts.

Condition-Specific Statistical Evaluation. To investigate whether the involvement of the network components was modulated by functional constraints of the linguistic input, we estimated condition-specific GC in the dominant connections extracted from the identified network components. The individual conditions were defined according to whether the words were presented in a well-formed sentence context (or were part of a word list), and according to whether the words were presented early in the sentence/word list (words two to four) or late in the sentence/word list ($n-3$ until $n-1$, with n being the number of words in the sentence/word list). To account for potential interpretational confounds of the resulting GC estimates we adopted a stratification procedure to ensure that, for each of the parcel pairs in each of the subjects, the marginal distributions of the epochwise signal variances as well as the words' lexical frequencies were equalized across conditions. Lexical frequencies were estimated using the Subtlex-NL database (crr.ugent.be/programs-data/subtitle-frequencies/subtlex-nl). Condition-specific histograms for lexical frequency were generated using 13 log-spaced bins. Histograms for signal variance were generated using six log-spaced bins. The consequence of this procedure is that only a subset of epochs is used for the subsequent estimation of GC, where the parcel-pair-specific number of epochs varies across parcel pairs. On average 50% of the epochs were retained (range: 20–75%), corresponding to 147 (range: 45–235) epochs.

From each of the extracted network components we defined a dominant connection as a spatially clustered set of edges that fulfilled the following criteria: (*i*) Each cluster consisted of at least four edges, (*ii*) the inflow/outflow nodes consisted of spatially adjacent cortical parcels, and (*iii*) nodes that for a given cluster of edges served both as input and output node were discarded, as well as the edges to which these nodes contributed.

This resulted in 42 connections for which we computed subject- and condition-specific GC, as an average across the contributing edges, and across the component-specific frequency range, defined by the IQR across subjects. We performed a nonparametric permutation test to evaluate the following contrasts: *i*) sentence – word list words, *ii*) for the sentence condition: early – late words, and *iii*) interaction effect: (early-late words sentences) – (early-late words sequences).

The statistical test performed was a two-sided permutation test (using 20,000 permutations) on Wilcoxon's signed rank statistic (Z-score) with a Bonferroni–Holm stepdown control for the familywise error rate. Fig. S2 shows in more detail data from the individual connections that showed statistically significant effects for one (or two) of the contrasts mentioned above.

Power Analysis. Our large sample of subjects provides a unique opportunity to perform a power analysis, to estimate the minimum sample size required to reliably identify the beta-band fronto-temporal and alpha-band temporo-frontal interaction patterns.

We randomly sampled from our pool of 102 subjects, using two different approaches. These approaches differed in the extent of reanalysis upon random sampling from our pool of 102 subjects. Our general approach was as follows. We randomly drew from the pool of subjects subsets of 10, 20, 30, or 40 subjects, and for each of these subsets we did this random selection 100 times. Next, we performed the NMF on the reduced-number-of-subjects concatenated GC spectra. In the first approach, we included only those edges that survived the statistical pruning procedure (based on a Bonferroni corrected parametric statistical evaluation of GC vs. reverse-time GC, as described in the main text), where this procedure was performed on the subselection of the subjects. In the second approach we included the 713 edges that were determined based on the statistical evaluation of the full batch of subjects. Next, we paired the NMF components obtained from the subsampled data with the 102-subject results, by computing pairwise correlations between the spatial topology of the components. The subsampled components with the highest spatial correlation with the original components were subsequently used. We focused on the temporo-frontal and fronto-temporal components, corresponding to the components in Fig. 2 *B* and *G* of the main text, respectively. Next, we defined a frontotemporal directionality index:

$$di = (FT_l - TF_l) + (FT_r - TF_r),$$

where FT_l is the average of the nonzero left-hemispheric fronto-temporal edges for a given component and TF_l is the average of the nonzero left-hemispheric temporo-frontal edges. We used Monte Carlo resampling to estimate the expected value of di under random allocation of the edge weights and subtracted the average across 500 Monte Carlo estimates from the di estimated from the observed component to obtain a corrected value. Thus, a di value >0 would indicate that the corresponding component is a predominantly frontal to temporal connection, and the a di value <0 would indicate a temporal to frontal connection.

Correct detection performance was calculated as the percentage of random subsets yielding a di value with the expected sign. The results are shown in the Table S2. When using the previously determined subset of 713 edges the detection performance was extremely good, even at low subject numbers, suggesting that only a moderate sample size is needed for a reliable detection of fronto-temporal beta-band and temporo-frontal alpha-band interactions. However, when using a more conservative approach, and using a data-driven approach to statistically prune the edges (using Bonferroni correction), based on the subset of subjects in the sample, detection performance became rather poor. This performance decrease was primarily caused by the conservative edge detection process's not selecting sufficient fronto-temporal and temporo-frontal edges to begin with. Extending the number of subjects in

the subselection to 50 and 60 increased the detection performance, but not in a satisfactory way. Relaxation of the statistical selection threshold (arbitrarily setting the edge acceptance threshold to 0.0005 uncorrected) recovered the detection rate to beyond 80% with a sample size between 30 and 40. Thus, based on these analyses we conclude that using a data-driven edge selection procedure with a statistical selection criterion that is not too strict would require a sample size of 30–40 subjects to detect the fronto-temporal and temporo-frontal connections with a probability of at least 80%.

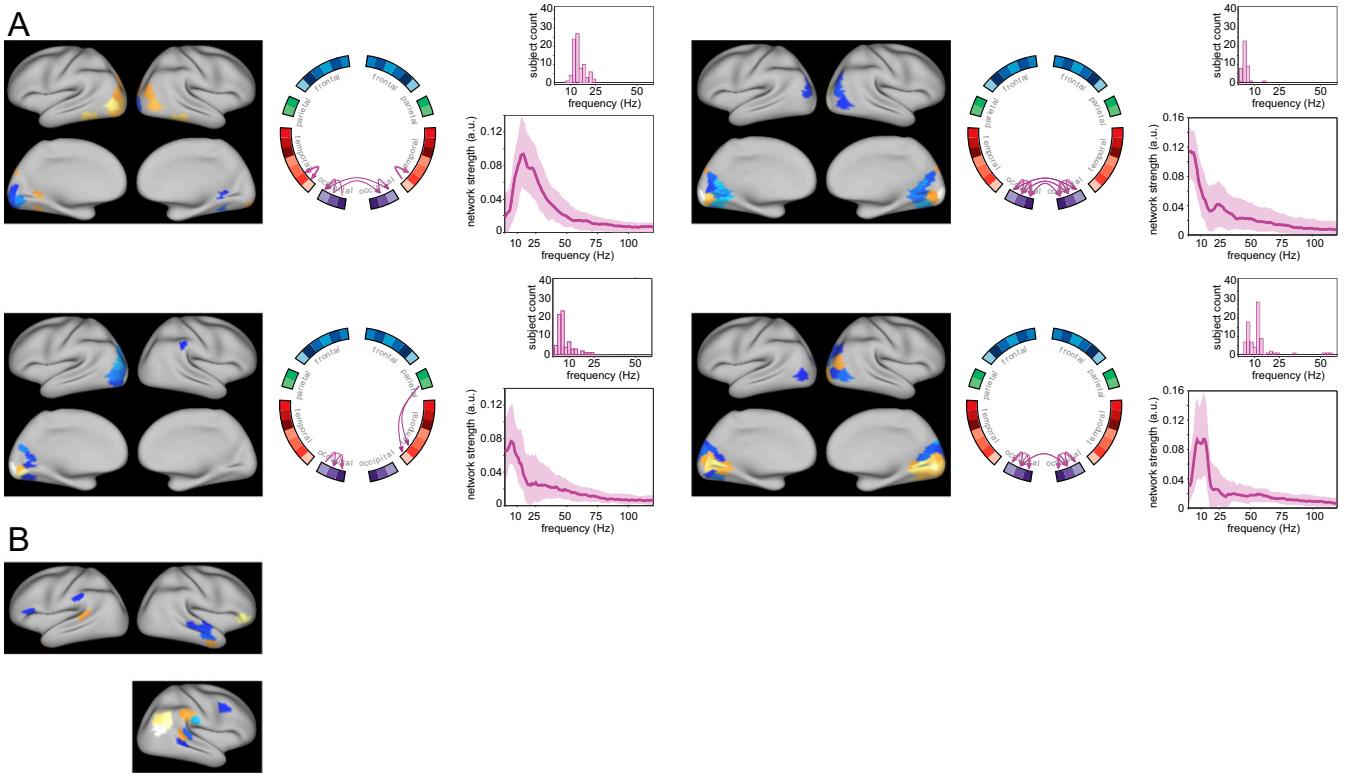
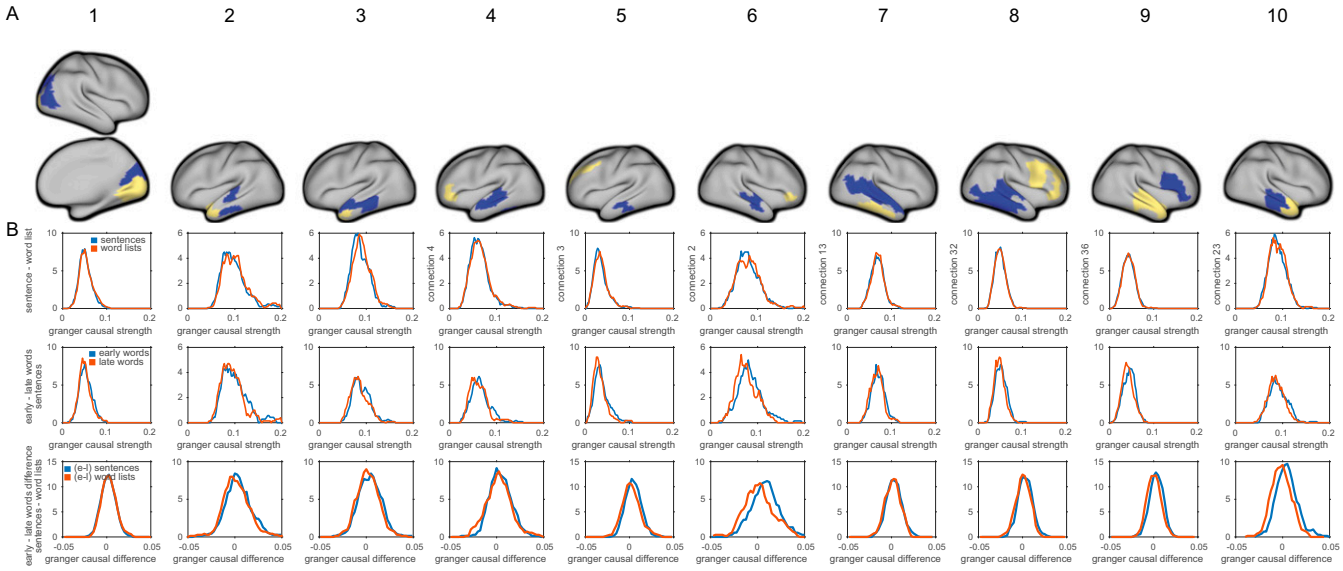


Fig. S1. Network components obtained by group-level NMF with predominant connections between visual cortical areas (*A*) and with spatially diffuse connections (*B*).



C

	Montecarlo p-values (uncorrected)			Frequency range (Hz)
	sentence - word lists	early words - late words	(e-l) sentence - (e-l) word lists	
1	0.00020	0.0041	0.29	10-14
2	0.00005	0.0015	0.10	4-13
3	0.00025	0.0004	0.02	14-19
4	0.018	0.0004	0.15	4-13
5	0.042	0.00005	0.026	4-13
6	0.094	0.00005	0.0008	4-13
7	0.042	0.00005	0.076	10-14
8	0.38	0.00005	0.0034	8-15
9	0.011	0.00005	0.0003	17-23
10	0.006	0.00005	0.00005	14-19

Fig. S2. Connection-specific data for the condition comparisons. (*A*) Spatial location of outgoing (blue) and ingoing (yellow) parcels. (*B*) Distribution of data values across subjects of Granger causal strength for the different conditions compared. Upper row: sentences vs. word lists. Middle row: early words vs. late words. Bottom row: difference values (early words minus late words) for sentences and word lists. (*C*) Table with uncorrected *P* values for the condition-specific comparisons, including a specification of the connection-specific frequency range.

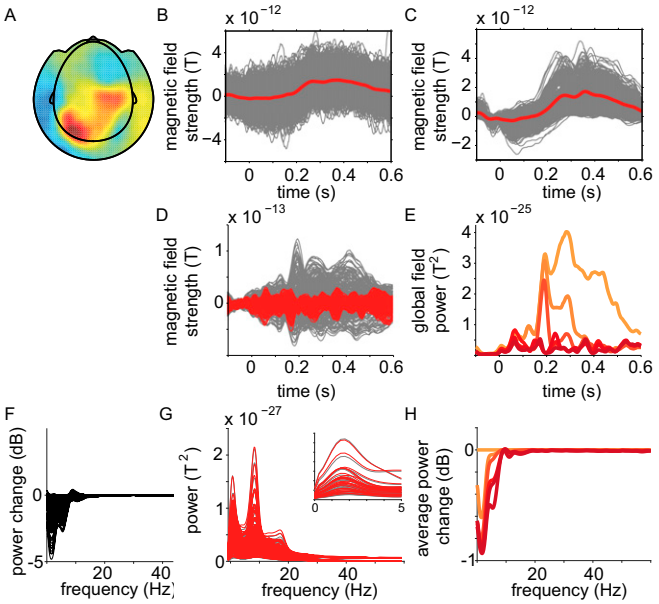


Fig. S3. Illustration of combined DSS/ASEO procedure tor the removal of word-onset event-related signal transients. (*A*) Spatial topography of the mixing coefficients for the first extracted DSS component for an example subject. (*B*) Single-trial time courses (in gray) of the first DSS component, time-locked to word onset, average across trials in red. (*C*) Single-trial estimates of the stimulus-locked transient response, estimated with the ASEO algorithm, average across trials in red. (*D*) Overlay of single channel event-related averages before (gray) and after (red) the cleaning procedure. (*E*) Global field power across channels of the event-related average (light orange) and after iterative removal of five DSS components (ranging from orange to dark red). (*F–H*) Effect of the DSS/ASEO procedure on spectral quantities. (*F*) Overlay of single-channel relative change in power after removal of event-related activity. (*G*) Overlay of single-channel power spectra before (gray) and after (red) the cleaning procedure. The magnified subpanel zooms in on a subset of channels for the low-frequency range. (*H*) Reduction in power (average across channels) during iterative removal of five DSS components.

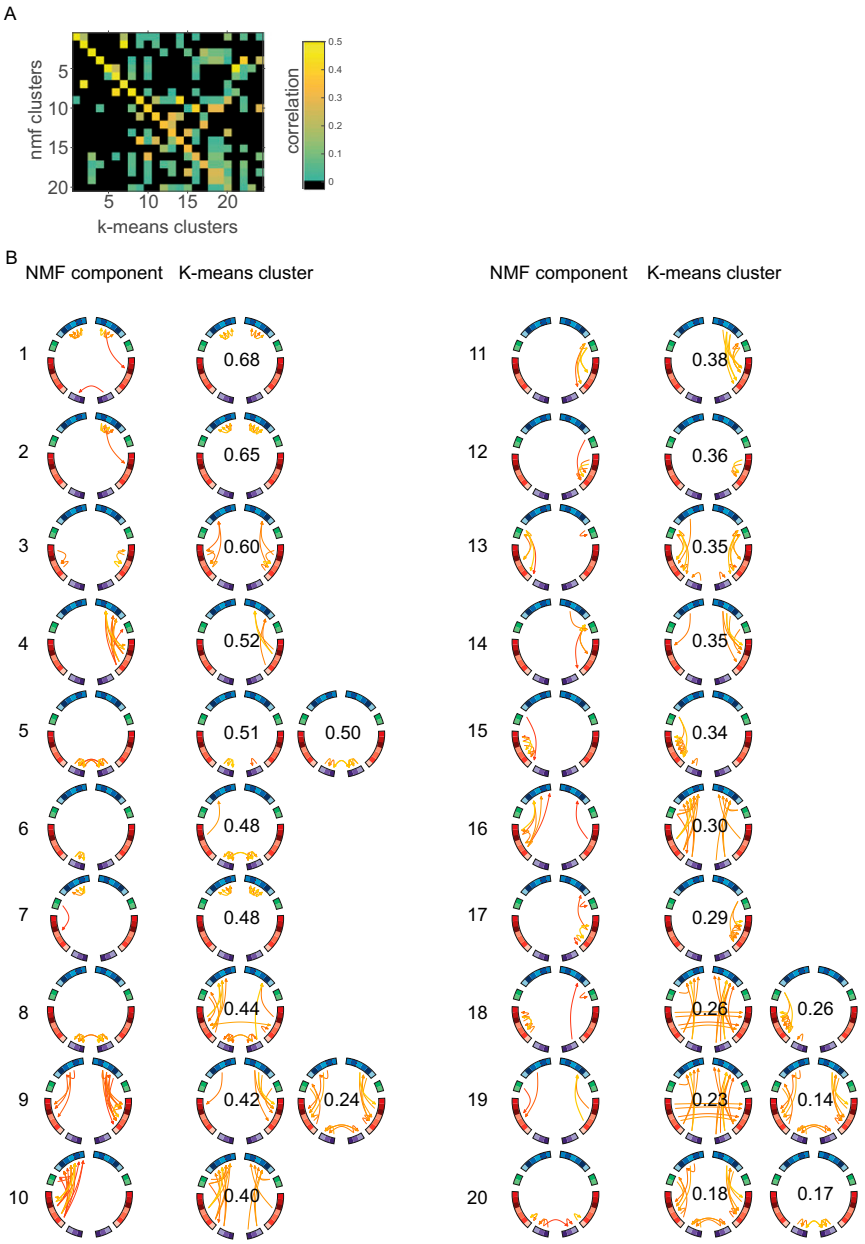


Fig. S4. Comparison between NMF and K-means clustering. (*A*) Pairwise correlation matrix between the spatial topology of network components identified with NMF and with K-means clustering. The components are ordered such that for each of the NMF components the most highly correlated K-means cluster is on the diagonal. (*B*) Spatial topology of NMF component (left columns) with corresponding highly correlated K-means cluster (right columns).

Table S1. Correlation between node degree and univariate signal properties

	Node degree inflow		Node degree outflow	
Signal property	Correlation	*P* value	Correlation	*P* value
Signal variance	0.054	0.51	0.077	0.34
Spatial filter norm	0.096	0.23	0.12	0.15
SNR	−0.097	0.23	0.032	0.69

Table S2. Detection probability of beta-band fronto-temporal and alpha-band temporo-frontal connections, based on subsampling of the subjects

No. of subjects	Fronto-temporal, %	Temporo-frontal 1, %	Temporo-frontal 2, %
Fixed set of edges (713), based on 102 subjects' results			
10	100	92	100
20	100	95	96
30	100	99	99
40	100	99	96
Data-driven edge detection (Bonferroni corrected *P* value)			
10	0	0	0
20	0	1	1
30	0.1	0.4	0.2
40	9	0.8	5
50	47	17	40
60	68	21	51
Data-driven edge detection ($P < 0.0005$, uncorrected)			
10	43		
20	90	41	35
30	90	77	77
40	95	91	87
50	96	92	93
60	95	94	91