NeuroImage 47 (2009) 1982–1991

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Neural network of speech monitoring overlaps with overt speech production and comprehension networks: A sequential spatial and temporal ICA study

Vincent van de Ven [a,*], Fabrizio Esposito [a,b], Ingrid K. Christoffels [c]

[a] Department of Cognitive Neuroscience, Faculty of Psychology, Maastricht University, PO Box 616, 6200 MD, Maastricht, The Netherlands
[b] Department of Neuroscience, University of Naples "Federico II," 80131 Naples, Italy
[c] Leiden Institute of Brain and Cognition (LIBC) and Leiden Institute of Psychology, Leiden University, Leiden, The Netherlands

ARTICLE INFO

Article history:
Received 20 November 2008
Revised 4 May 2009
Accepted 18 May 2009
Available online 27 May 2009

ABSTRACT

The neural correlates of speech monitoring overlap with neural correlates of speech comprehension and production. However, it is unclear how these correlates are organized within functional connectivity networks, and how these networks interact to subserve speech monitoring. We applied spatial and temporal independent component analysis (sICA and tICA) to a functional magnetic resonance imaging (fMRI) experiment involving overt speech production, comprehension and monitoring. SICA and tICA respectively decompose fMRI data into spatial and temporal components that can be interpreted as distributed estimates of functional connectivity and concurrent temporal dynamics in one or more regions of fMRI activity. Using sICA we found multiple connectivity components that were associated with speech perception (auditory and left fronto-temporal components) and production (bilateral central sulcus and default-mode components), but not with speech monitoring. In order to further investigate if speech monitoring could be mapped in the auditory cortex as a unique temporal process, we applied tICA to voxels of the sICA auditory component. Amongst the temporal components we found a single, unique component that matched the speech monitoring temporal pattern. We used this temporal component as a new predictor for whole-brain activity and found that it correlated positively with bilateral auditory cortex, and negatively with the supplementary motor area (SMA). Psychophysiological interaction analysis of task and activity in bilateral auditory cortex and SMA showed that functional connectivity changed with task conditions. These results suggest that speech monitoring entails a dynamic coupling between different functional networks. Furthermore, we demonstrate that overt speech comprises multiple networks that are associated with specific speech-related processes. We conclude that the sequential combination of sICA and tICA is a powerful approach for the analysis of complex, overt speech tasks.

Introduction

Speakers continuously monitor their speech during production, which suggests that speech monitoring is an integral part of the speech production process (Levelt et al., 1999; Postma, 2000). During speaking the speaker is effectively presented with auditory feedback of the overt speech, which may be processed by the speech comprehension system to serve the monitoring process. Evidence for this concept of speech monitoring comes from studies in which speech feedback was absent or distorted during speaking, which resulted in impaired monitoring and control of speech (Fu et al., 2006; Lane and Tranel, 1971; Postma and Kolk, 1992). Thus, speech production, comprehension and monitoring are strongly intertwined processes that could be largely underpinned by overlapping neural networks. However, it remains unclear to what degree speech monitoring is embedded within neural substrates of speech production and comprehension. In this study, we investigated how spatially distributed speech-related brain functional networks contributed to different speech tasks.

Several neuroimaging studies showed that monitoring of one's own speech involves areas of the superior temporal gyrus (STG) (e.g., Hashimoto and Sakai, 2003; McGuire et al., 1996). These studies manipulated speech monitoring by altering the overt speech feedback to the participants, for example, by altering the pitch of on-line speech feedback (McGuire et al., 1996; Toyomura et al., 2007), or by delaying the speech feedback with respect to the moment of speaking (Hashimoto and Sakai, 2003). Altered feedback made the labeling of the perceived speech as one's own more difficult (Fu et al., 2006; McGuire et al., 1996). This effect was associated with increased hemodynamic responses during altered feedback conditions, in comparison to unaltered speech feedback. In addition, a recent study showed that brain activity during the on-line monitoring of one's own (unaltered) speech feedback is decreased compared to masked feedback, as well as passive speech perception (Christoffels et al., 2007). These results correspond to theoretical models that

* Corresponding author.
E-mail address: v.vandeven@psychology.unimaas.nl (V. van de Ven).

1053-8119/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2009.05.057

suggest that sensory areas may be inhibited during motor acts in order to cancel out the resulting self-generated sensory effects (e.g., Blakemore et al., 1998; Paus et al., 1996; Wolpert et al., 1995). More specifically, these 'forward models' of sensory inhibition by motor acts predict that the amount of overlap between the expected consequences of articulation and actual auditory feedback determines the amount of attenuation of sensory cortex activation. Several neurophysiological findings in the monkey (Eliades and Wang, 2003, 2005), electroencephalography (Ford et al., 2002) and microelectric recordings in humans (Creutzfeldt et al., 1989) provide empirical support for the predictions of the forward models.

In addition, monitoring or labeling of perceived speech as one's own may also be associated with medial frontal and parietal areas that have been associated with a default-mode network (DMN) of brain activity (Raichle et al., 2001). The DMN has recently received increased interest, because it may be associated with an intrinsic organization of functional networks in the brain. The DMN commonly shows deactivation during task performance, in comparison to a baseline. A few studies reported changes in activity in the posterior cingulate cortex and precuneus during speech processing (Binder et al., 1997; Jardri et al., 2007), in which these areas may contribute more to self-generated, internal speech than to passive listening (Jardri et al., 2007). In addition, default-mode areas have been associated with perspective-taking (Lindner et al., 2008) and self-referential processing (Beer, 2007). Thus, with respect to speech processing, the DMN may be more associated with agency and self-referential aspects of speech, rather than the passive perception of speech. Moreover, the degree of deactivation may vary with the extent to which speech monitoring is recruited during on-line, overt speech production.

The neural correlates of on-line speech monitoring have not been investigated often on a system level (cf, Tourville et al., 2008). Such an analysis may not only reveal which areas form a functional network that is associated with speech monitoring, but may also reveal if and how such a network is embedded within functional networks of speech production and comprehension. To address these questions, we investigated the functional brain connectivity during an overt speech production and monitoring task of a previously published functional magnetic resonance imaging (fMRI) study (Christoffels et al., 2007). More specifically, we attempted to isolate the speech monitoring network from other speech-related processing networks. We used spatial independent component analysis (sICA) (McKeown et al., 1998) to decompose whole-brain blood oxygenation level-dependent (BOLD) signals into spatial networks, and investigated the temporal contribution of these components to the functional data. Although the cognitive subtraction method (typically a mass-univariate, voxel-by-voxel analysis) has been quite successful in localizing task-related activity of a number of core regions involved in speech monitoring (Christoffels et al., 2007; Toyomura et al., 2007), it is not suitable to study brain dynamics of speech monitoring on a network level, which requires a multivariate analysis to quantify relations between voxel time courses. In contrast, ICA provides a way to investigate how different networks support different aspects of the complex task of speech production. In general, ICA is a multivariate and data-driven method that assumes that the observed data is a linearly weighted combination of a set of statistically independent sources (Comon, 1994). ICA decomposes a dataset into components (source estimates) by maximizing their mutual statistical independence. In practice, both the sources and the weighting parameters are unknown and need to be estimated from the observed data. In fMRI research mostly spatial ICA has been used, because typically the number of voxels is much larger than the number of scans (Calhoun et al., 2001; McKeown et al., 1998). This means that components are characterized by maximized independence in the spatial domain. As a consequence, sICA provides estimates of new basis functions (component time courses of activity) that are not subjected to the independence criterion, but are left free to correlate. Thus, sICA is able to estimate multiple spatial networks of functional connectivity in a single analysis, without the requirement of a priori specification of an ideal hemodynamic signal.

Our spatial ICA results showed that several distinct networks of cortical and subcortical areas contributed in unique ways to the processing of the overt speech tasks (see below). However, we did not obtain a spatial component that exclusively showed a temporal profile characteristic of speech monitoring, that is, attenuated response during speech with normal feedback (Christoffels et al., 2007). One explanation why this was the case may be that the spatial independence criterion was not optimal to identify the speech monitoring signal. We therefore applied temporal ICA (tICA) (Biswal and Ulmer, 1999; Calhoun et al., 2001; Seifritz et al., 2002) to a restricted part of the data by selecting voxels of the spatial auditory component. This procedure revealed a single and unique temporal component that exclusively represented the speech monitoring temporal pattern sensu Christoffels et al. (2007). Further confirmation that the temporal component was associated with speech monitoring was obtained by regression analysis of the temporal component to the whole-brain functional imaging data, which showed the strongest contributions of bilateral Heschl's sulcus to the temporal component. Furthermore, the temporal component was strongly negatively correlated with activity in the supplementary motor area (SMA). Consecutive post-hoc psychophysiological interaction analysis (PPI) (Friston et al., 1997; Kim and Horwitz, 2008) showed that functional connectivity between SMA and bilateral Heschl's sulcus changed with task conditions. These findings demonstrated that the temporal component showed a coupling between different spatial networks, which changed with task conditions.

In the following sections, we first describe the methods and results of spatial ICA of the functional time series, and then describe the methods and results of temporal ICA of the spatially restricted set of time series. The interpretations and inferences of the sICA and tICA results are considered in the discussion section.

Methods

Participants

Data of twelve participants from a previous overt speech fMRI study was used (Christoffels et al., 2007). The mean (SD) age of the participants was 23.4 (1.4), and all were right handed. All participants were native Dutch speakers and had no history of hearing or language related problems. All participants gave their written informed consent before participating in the study. The study was approved by the ethical committee of the University Medical Center of Nijmegen, The Netherlands.

Experimental procedure

The experimental procedure has been described previously (Christoffels et al., 2007) and is briefly reported here. In a blocked design, five different task conditions were presented, alternating with a fixation condition: 1) Overt picture naming [*PNvoice*], 2) Overt picture naming with added pink noise [*PNnoise*], 3) Covert picture naming [*PNcovert*], 4) Listening to previously recorded own speech [*LISvoice*] and 5) Listening to pink noise [*LISnoise*]. In the overt picture naming conditions participants were required to name presented pictures as quickly and accurately as possible. However, in PNnoise participants heard additional loud pink noise that masked the perception of their speech. In the covert naming condition participants covertly generated the picture name. In the two passive listening conditions the auditory input was similar to the picture naming conditions: Both the participants' voice (LISvoice) and the pink noise (LISnoise) were presented, but participants were not required to give a naming response.

In all conditions white-on-black line drawings were presented for 1000 ms, with an onset of 100 ms after the beginning of the silent delay between volume acquisitions. There was 2250 ms of silence before the next volume was acquired in which the participants could respond. A fixation cross was presented immediately before picture onset and during fixation conditions. During the listening conditions scrambled pictures were presented and the auditory stimulus was presented using the same timing as the PNnoise condition. For the LISnoise condition the same pink noise recording was used as in the PNnoise condition.

Each experimental block consisted of five trials of one condition (20 s). The fixation blocks (16 s) consisted of a continuous presentation of a fixation cross and a brief instruction (2 s) about the condition of the upcoming experimental block. There were five functional runs of 15 experimental task blocks that alternated with fixation blocks (i.e., 3 repetitions of each condition per run, in total 15 repetitions of each condition per participant).

Imaging parameters

Imaging was performed on a 3 T whole-body system (Magnetom Trio, Siemens Medical Systems, Erlangen, Germany) using a standard head coil. Functional volumes were acquired using a T2*-weighted echoplanar sequence with BOLD contrast (27 slices; TR/TE, 4000/30 ms; volume scanning time, 1750 ms; inter-scan gap, 2250 ms; FA, 90°; field of view, 224 mm^2; slice thickness, 4.5 mm (no slice gap); matrix, 64×64; voxel size, $3.5\times3.5\times4.5$ mm^3). A total of 705 volumes were acquired for each participant in 5 runs of 141 volumes each. High-resolution anatomical volumes were acquired using a T1-weighted three-dimensional (3-D) MP-RAGE (Magnetization-Prepared Rapid Acquisition Gradient Echo) sequence (192 sagittal slices; TR/TE, 2300/3.93 ms).

Data analysis

Anatomical and functional images were pre-processed and co-registered using BrainVoyager QX (Brain Innovation, Maastricht, The Netherlands). The first two volumes of every run were discarded to account for the T1 saturation effect. Functional image pre-processing steps included slice scan time correction, linear trend removal and temporal high-pass filtering (3 cycles per time course $\approx$ 0.0054 Hz). Pre-processed functional time series were then transformed into 3-D standardized space (Talairach and Tournoux, 1988) using a resampled voxel size of $3\times3\times3$ mm^3. Anatomical data were resampled and transformed to the same standardized space using a resampled voxel size of 1 mm^3. A volume mask was created from the average of the anatomical 3-D images of the participants, which excluded voxels belonging to the ventricles or that were outside of the average brain. Voxel time courses that were tagged by the mask were used for further analysis using Matlab routines (The Mathworks Inc, Mass., USA).

Spatial ICA

The functional data were analyzed using spatial ICA within a framework that we here refer to as hierarchical-ICA (hICA), which is an extension of a previously published self-organizing group-ICA methodology (sogICA) (Esposito et al., 2005) from single to multiple runs per subject. HICA was applied to the functional datasets in the following way (see Fig. 1). First, each masked dataset was individually decomposed by spatial ICA, using the Infomax gradient-learning algorithm (Bell and Sejnowski, 1995; McKeown et al., 1998) into a set of 35 components (learning rate = 0.0001; stopping criterion $\varepsilon = 0.0001$; batch size = 25; retained variance > 99%, dimension reduction performed using principal component analysis [PCA]). After decomposition the spatial components were *Z*-scored

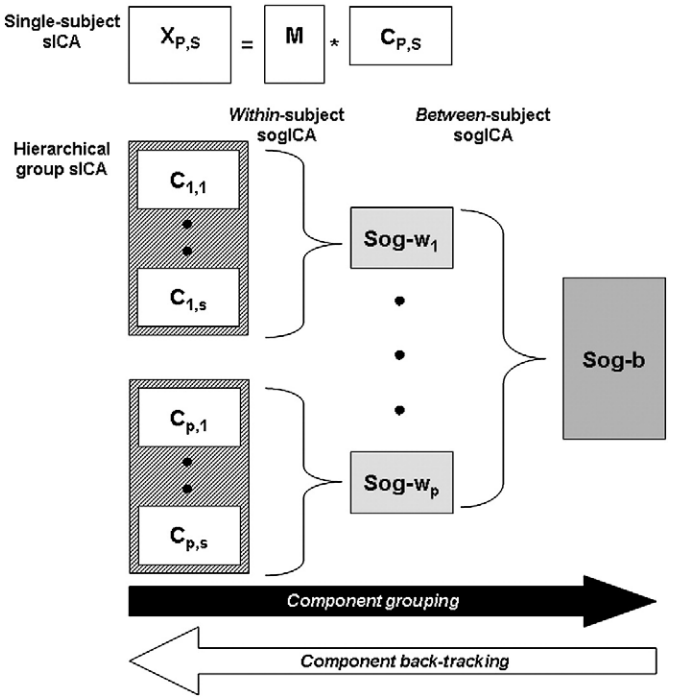

Fig. 1. Processing steps of hierarchical group-ICA (HICA). Upper row: The spatial ICA model for a single-subject decomposition (McKeown et al., 1998). HICA groups components of individual decompositions (hatched boxes) into within-subject clusters (step 1, left-most accolades, resulting in light grey boxes), which are then clustered into between-subject clusters (step 2, right most accolade, resulting in dark grey box). Each step results in a set of component clusters, where the original source of each component can be retrieved by back-tracking the clustering process. X, time course data matrix of size voxels × timepoints; M, mixing matrix; C, component matrix of size components × timepoints; P, participant 1 .. p; S, session 1 .. s; Sog-w(b), self-organizing grouping result of the within- (between-) subject clustering.

(McKeown et al., 1998). Then, sogICA was used to cluster the decompositions of the *within-subject* repetitions of the functional runs using spatial correlation as similarity measure, which resulted in 35 clusters of 5 components each for each subject. For each within-subject cluster an average cluster group map was obtained, which served as input to the second-level *between-subject* clustering, which yielded 35 clusters of 12 components each. Between-subject group components were calculated as a one-sample *t*-statistic map against a fixed mean of 0. Spatial templates of target cortical areas were used to select the best matching final group components for further investigation (Van de Ven et al., 2004; van de Ven et al., 2008), which included 1) bilateral auditory cortex, 2) left inferior frontal gyrus and superior temporal cortex, 3) pre- and postcentral gyri and 4) posterior cingulate cortex. In addition, we also used templates for 1) striate and 2) extrastriate cortex and 3) bilateral frontal and parietal cortices in order to identify components that were associated with the visual aspects of the task (task instructions and picture presentation). The between-subject group maps were thresholded at $t(11) = 3.1$ (one-sided, $P \leq 0.005$, uncorrected), which were then each corrected at the voxel-cluster level at a false-positive rate of $\alpha = 0.05$ (Forman et al., 1995; Goebel et al., 2006). This procedure provides a multiple comparison correction at the voxel-cluster level and estimates a corrected threshold for voxel-cluster size, given the inherent spatial smoothness of the statistical map. The spatial smoothness of the initial (uncorrected) statistical map was used to reiteratively simulate 1000 maps (Monte Carlo simulation). Of each simulated map, clusters of voxels that surpassed the uncorrected voxel-level threshold were gathered in a list. The gathered list of voxel-cluster sizes was then thresholded at a false-positive rate of 5%, and the minimum cluster size at that threshold was used as the new cluster size threshold on top of the initial voxel-level threshold (see also Van de Ven et al., in

press). Thresholded maps were superimposed on inflated and flattened anatomical representations (courtesy of Montreal Neurological Institute [MNI]).

The temporal characteristics of the group components were analyzed as follows. By *back-tracking* the hICA procedure for a selected group component the corresponding spatial component for each dataset was identified (see Fig. 1). Then, for each component the component-activity time course was calculated by *back-projecting* the component map to the original data space (Duann et al., 2002; Esposito et al., 2003; McKeown et al., 1998), and standardized to units of percent signal change. Time courses were then averaged within participants, according to the onsets of the experimental conditions and finally averaged across participants. Amplitude changes were calculated from the event-related averages by subtracting time course values of the baseline (selecting timepoints −1 to 0 prior to onset of conditions) from values during peak responses (timepoints 2 to 4 after condition onset) for each participant. Significance of the amplitude changes from baseline was calculated by one-sample *t*-tests against a fixed mean of 0.

It should be noted that component maps and time courses were inspected at the end of the grouping pipeline. Thus, component clustering was done in a blind and automated fashion where the final results were used to select groups of components for further analysis.

Results

Selection of hICA group components was based on best match with a set of spatial templates, which resulted in seven clusters of interest. Four of the clusters comprised time courses that showed condition-specific changes in activity: 1) Auditory cluster (AC), 2) Central sulcus and motor cluster (CS), 3) Left fronto-temporal cluster (lFT) and 4) Default mode network cluster (DMN). Three other group components comprised time courses that showed condition-unspecific changes in activity, which were related to the visual aspects of the task: 1) (primary) visual cortex cluster (PVC), 2) dorsal and ventral extra-striate cortex cluster (EVC) and bilateral fronto-parietal cluster (bFP). In the following, the condition-specific clusters are described. The condition-unspecific clusters are described in the Supplementary results.

Condition-specific components

Figs. 2A, B show the four condition-specific clusters superimposed on inflated representations of a standardized anatomical image. Table 1 lists the coordinates of the voxel-clusters of the four clusters.

The AC component (red–yellow color coding in Fig. 2; spatial map corrected at cluster-size threshold of 286 mm^3, $P<0.05$) comprised several areas of the STG including Heschl's gyrus and sulcus, the middle temporal gyrus (MTG), and insula. The cluster showed increased activity for those conditions that provided speech-related auditory input (*PNvoice*: $t(11)=2.4$, $P=0.037$; *PNnoise*: $t(11)=3.2$, $P<0.01$; *LISvoice*: $t(11)=2.2$, $P=0.047$), suggesting that it represented sensory auditory processing.

The CS component (blue color scale; cluster threshold = 388 mm^3, $P<0.05$) comprised the bilateral central sulcus, bilateral precentral gyrus, SMA and several subcortical areas including thalamus and striatum. The component showed increased activity specifically for overt speech production (*PNvoice*: $t(11)=5.0$, $P<0.001$; *PNnoise*: $t(11)=5.2$, $P<0.001$), and to a lesser extent covert speech production

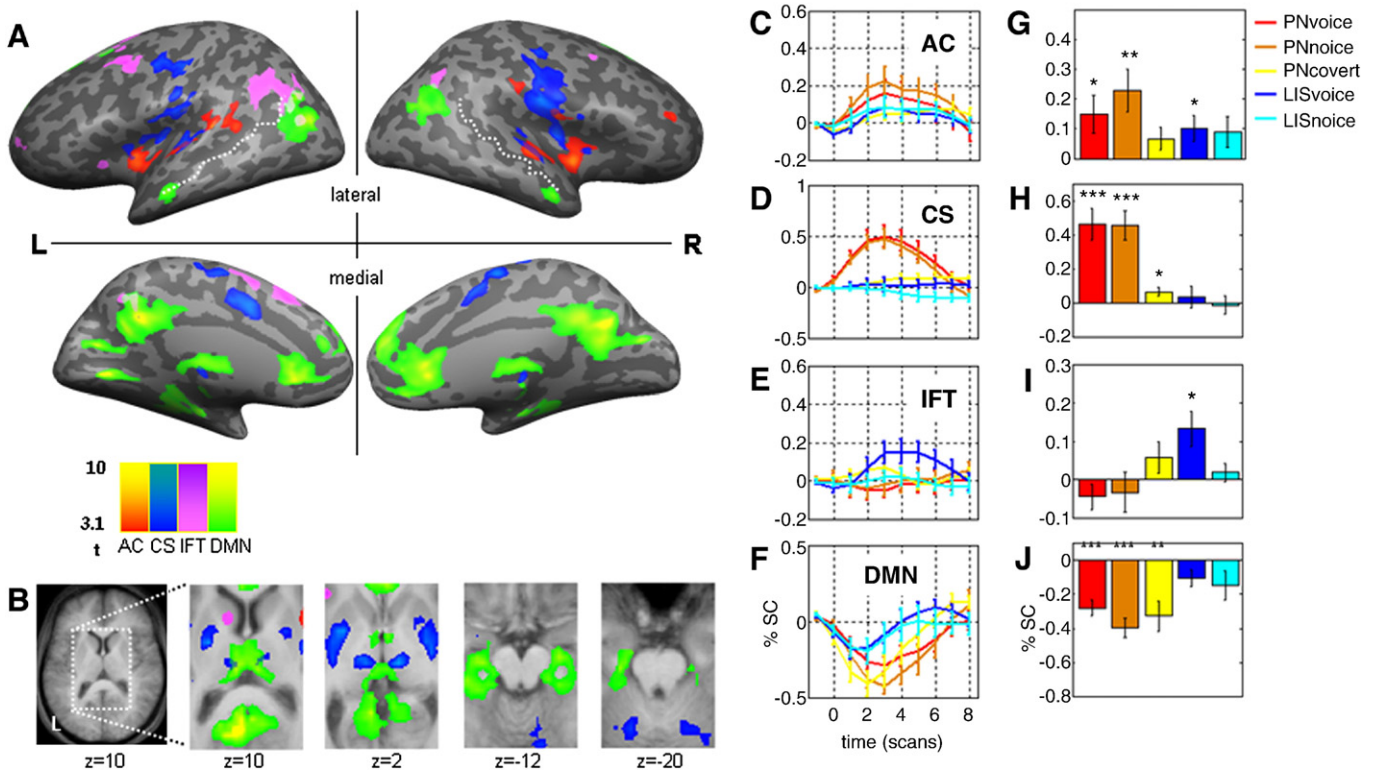

Fig. 2. Condition-specific HICA components. A: Spatial group maps of the auditory cluster (AC; red–yellow color coding), central sulcus cluster (CS; blue–light blue), left fronto-temporal cluster (lFT; pink–purple) and default-mode network cluster (DMN; green–yellow) are superimposed on left and right inflated hemispheres, thresholded at $t(11)=3.1$ (one-sided, $P=0.005$). White dotted line delineates the superior temporal sulcus. Panels in B show zoomed-in subcortical structures (see left-most panel) on which the group maps are superimposed. The CS cluster includes caudate and thalamic nuclei, and cerebellar areas. The lFT cluster includes the left putamen. The DMN cluster includes medial thalamic and hippocampal areas. Left hemisphere is shown on the left of each panel. Event-related averages (C–F) and bar plots (G–J) of the back-projected component time courses are shown for AC (C,G), CS (D,H), lFT (E,I) and DMN (F,J) (time-locked to block onset). The abscissa of the time course plots refers to the number of scans, the ordinate refers to percent signal change (% sc). Note that each spatial component has a unique, condition-specific contribution to the data. Error bars indicate 1 standard error; *$P<0.05$; **$P<0.01$; ***$P<0.001$.

Table 1
Condition-specific components of interest (COIs).

COI	x	y	z	Size	t max	Mean p	Area
bAC	−52	−31	14	4642	6.3	0.0027	L HS/STG
	51	−11	6	2803	5.4	0.0022	R HG/HS
	36	−1	1	2033	9.6	0.0022	R Ins
	−38	−6	0	1876	8.0	0.0019	L Ins
	42	38	15	876	6.4	0.0026	R MFG
	52	−25	23	516	5.0	0.0029	R PO
	−57	−51	−2	434	5.1	0.0030	L MTG
	−2	−55	46	361	4.9	0.0029	PreC
CS	44	−9	15	10,727	6.9	0.0021	R CS/Ins/PO
	0	−5	47	7784	9.0	0.0014	SMA/aCC
	−48	−17	21	7703	6.5	0.0022	L CS/Ins/PO
	−25	−6	2	3070	8.6	0.0014	L Putamen
	−10	−18	2	1480	6.6	0.0015	L Thal
	10	−18	1	1093	8.6	0.0015	R Thal
	14	−58	−17	1014	5.1	0.0030	R Cerebellum
	−16	−55	−20	476	4.6	0.0031	L Cerebellum
lFT	−16	13	48	9881	10.9	0.0020	L SFG
	−44	−61	32	4828	5.4	0.0024	L AG/IPS
	−3	−57	37	1573	6.0	0.0022	PreC
	21	−69	−28	1250	4.7	0.0030	R Cerebellum
	−13	2	14	519	6.8	0.0020	L Caudate N
	−56	−28	−8	506	4.7	0.0029	L MTG
	43	−56	36	470	4.7	0.0025	R AG
	−33	19	0	387	7.0	0.0021	L Ins
DMN	−2	−43	14	40,439	17.9	0.0011	pCC/Hipp/parahippo
	0	38	12	19,385	12.1	0.0010	aCC
	−43	−68	21	3790	8.2	0.0018	L IPC/MTG
	44	−58	23	2820	6.4	0.0017	L IPC/MTG
	0	−42	−39	1454	6.9	0.0019	Cerebellum
	21	23	48	1432	5.7	0.0024	R SFG
	−55	−14	−11	1285	7.7	0.0019	L MTG
	−18	23	46	1106	6.1	0.0024	L SFG
	54	−7	−14	1015	6.5	0.0022	R MTG

Talairach coordinates (*x,y,z*) are reported for the center of mass of voxel clusters. *t*-values represent maximum value of the cluster ($df = 11$). HS(G), Heschl's sulcus (gyrus); STG, superior temporal gyrus; MTG, middle temporal gyrus; Ins, Insula; SFG, superior frontal gyrus; MFG, middle frontal gyrus; AG, angular gyrus; a(p)CC, anterior (posterior) cingulate cortex; PreC, Precuneus; PO, parietal operculum; Thal, Thalamus; Hippo, Hippocampus; parahippo, parahippocampal gyrus; IPC, inferior parietal cortex.

(*PNcovert*: $t(11) = 2.4$, $P = 0.033$). We suggest that this component represented the involvement of facial motor processes of speech production.

The lFT component (pink–purple color; cluster threshold = 348 mm^3, $P<0.05$) comprised largely left-lateralized voxel-clusters in inferior, middle and superior frontal gyri, inferior parietal and superior temporal gyri, and left striatal nuclei. This component was mostly activated during listening to prerecorded speech (*LISvoice*: $t(11) = 2.8$, $P = 0.017$), suggesting strong sensitivity to higher-order speech comprehension processes.

The DMN component (green–yellow color; cluster threshold = 746 mm^3, $P<0.05$) included the posterior cingulate cortex, the anterior cingulate cortex, bilateral inferior parietal, anterior regions of the middle temporal cortex, thalamus, and hippocampus and parahippocampal areas. This cluster showed the strongest deactivations for overt and covert speech production (*PNvoice*: $t(11) = -6.1$, $P<0.001$; *PNnoise*: $t(11) = -6.8$, $P<0.001$; *PNcovert*: $t(11) = -3.8$, $P<0.01$). We suggest that this component is mostly associated with the on-line processing of self-referential aspects of speech.

Supplementary Fig. S2 shows the spatial overlap and differences between sICA results and linear regression main effects of picture naming and listening conditions. For the main effects, white coloring refers to increased activity with respect to baseline and black coloring refers to decreased activity. Increased activity during picture naming (main effect) overlapped with the CS and AC component, and increased activity during listening (main effect) overlapped with the AC component. These findings of overlap would be expected given the AC and CS component time courses. The picture naming main effect also showed decreased activity in medial frontal cortex, which overlapped with the frontal areas of the DMN component. The listening main effect also showed increased activity in medial superior frontal areas, which overlapped with areas in the lFT component. However, main effects did not include lateral areas of the DMN and lFT components.

Temporal ICA

The spatial components showed clearly circumscribed connectivity patterns, and showed unique contributions to the overt speech conditions. However, no single component clearly showed a characteristic speech monitoring network in the sense of altered activity of normal feedback in comparison to manipulated feedback (Christoffels et al., 2007; Fu et al., 2006; McGuire et al., 1996) within the STG. In light of a previous analysis of this data, as well as reports of other studies, we hypothesized that the effect of discrepancies between expected and actual verbal feedback resided within the auditory areas (Christoffels et al., 2007; Tourville et al., 2008). To investigate if the neural correlate of speech monitoring could be represented by a temporally independent source that was spatially distributed across multiple networks we used temporal ICA (tICA).

The application of tICA to fMRI datasets is less straightforward than spatial ICA, mostly because the number of voxels (the channels in tICA) is far greater than the number of timepoints (samples). When left unrestricted, estimation of the covariance matrix may become unstable, and poses a computational limitation for most working stations. An often-used strategy is to apply tICA to a subset of voxels (Calhoun et al., 2001; Seifritz et al., 2002), usually in the context of a spatially-informed restriction, i.e., when temporal sources are hypothesized to be present in one or several specific brain structures (Seifritz et al., 2002). In our case, we used the AC spatial group component as a mask to select voxel time courses for further analysis (voxel selection criterion: $t \geq 3.1$).

A second issue that must be dealt with is the requirement of temporal independence within the restricted dataset. A rule-of-thumb is that the number of samples should be some multiple of the square of the number of channels (Makeig et al., 1999). To ensure that the restricted data met this criterion, the restricted volume of each of the five functional sessions were first standardized (via *Z*-transformation) and temporally smoothed (Gaussian kernel FWHM = 6.2 s) and then concatenated within each participant, resulting in an aggregate dataset of 492 voxels by 695 timepoints for each participant. We then restricted the dimension of each aggregate dataset to 25 (using PCA), which resulted in a reduced dataset that provided 25×695 inputs to estimate 25^2 unmixing weights.

Infomax ICA (Bell and Sejnowski, 1995) was used to decompose each participant's reduced aggregate dataset into 25 temporally independent components. The components were clustered across participants using sogICA (Esposito et al., 2005), but for this analysis the clustering was applied to the temporal dimension only (temporal correlation as similarity measure) (see Fig. 3A). The resulting group temporal components were ranked in ascending order of intra-cluster distance. The group temporal components were further investigated by creating condition-specific event-related averages for each subject using the standardized temporal components, and were then averaged across participants (thus, in a similar fashion as event-related averages of the spatial components). Amplitude changes were calculated in the same way as those calculated for the event-related averages of the spatial components. The event-related averages of the temporal component clusters were visually inspected for a speech monitoring effect sensu Christoffels et al. (2007). To quantify the effect, we used a repeated measures analysis of variance (ANOVA) to calculate a 2-way interaction of the amplitude changes between speech source (PN vs. LIS) and feedback

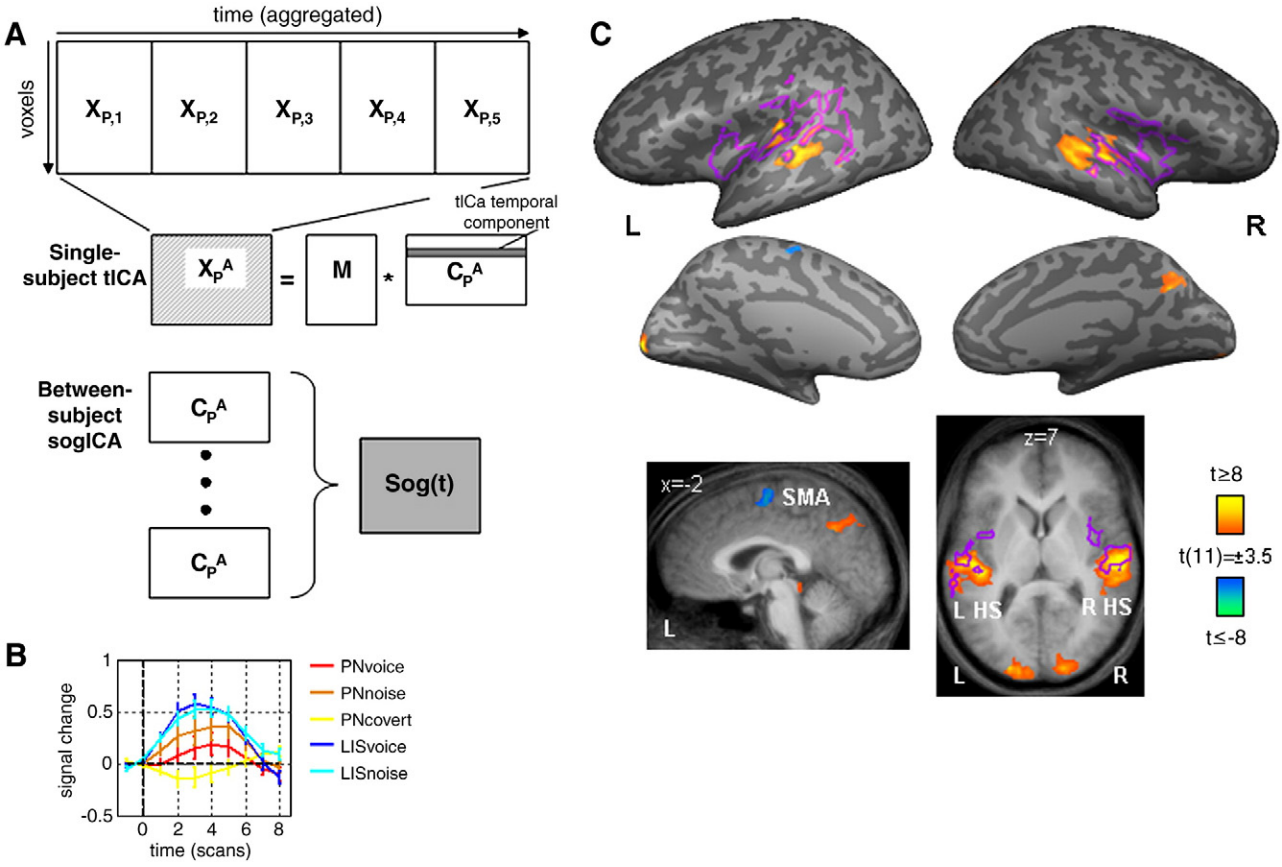

Fig. 3. Temporal ICA method and results. (A) Processing steps of single-subject and consecutive grouping of temporal ICA (tICA). Upper row: The tICA model for a single-subject decomposition, in which selected time courses of the five sessions of a participant are standardized (*Z*-scoring) and aggregated into dataset X^A. The aggregated dataset is then decomposed using tICA, which results in a component matrix (C) in which each row is a temporally independent component. Rows of M indicate the spatial contribution of the temporal components to the aggregate dataset. Clustering across individual realizations of C is performed in one analysis step, using sogICA, in which the similarity measure is temporal pair-wise correlations (compare with Fig. 1). Sog(t), self-organizing grouping result of the between-subject clustering of temporal components. Other abbreviations and notations are similar to Fig. 1. (B) Selected tICA component, which shows decreased activity for overt speech (PNvoice) compared to overt speech with noise (PNnoise), which are both decreased in comparison to the listening conditions (LISvoice and LISnoise). The ordinate refers to signal change in arbitrary (standardized) units. (C) The fit of the temporal component to the functional data is presented as a random effects fit with individual design matrices that were generated using a leave-one-out procedure. Results are superimposed on left and right inflated hemispheres (see Fig. 2A) and on the averaged anatomical image (map thresholded at $P<0.005$, two-tailed, and corrected at the cluster-level size of 104 mm^3, $P<0.05$; left hemisphere on the left of pictures). Positive fits included bilateral Heschl's sulcus (HS), precuneus and occipital cortices. Negative fits left bilateral supplementary motor area (SMA). Purple outlines represent the outline of voxel clusters of the AC spatial component that were used to select voxels for tICA. Voxel cluster sizes and coordinates are presented in Table 2.

type (voice vs. noise). To obtain a full-factorial design, we ignored the amplitude changes to the covert picture naming condition.

We also investigated the contribution of the selected cluster of temporal components to the whole-brain time courses. One option for this analysis could be to use the temporal components for each participant as the ideal model of brain activity for that participant. However, this entails a circularity in which the data of each participant is fitted to a basis function that was derived from that same data. To avoid this circularity, we adopted a leave-one-out procedure for creating individual GLM predictors. In this procedure, the design matrix for a particular participant was built by averaging the temporal components from all but one participant, leaving out the temporal component of the tested participant. Because the temporal components were estimated from individual participant data sets, the leave-one-out strategy ensured that no temporal component estimated from one data set was re-used to explain the variance of that same data set. This step was repeated twelve times to build a new design matrix for each participant. Finally, for each participant an additional standard predictor was added to the design matrix as second covariate, which modeled the presence of the task, irrespective of condition. Design matrices were fitted in a two-level, random effects analysis. Results were initially thresholded at $P\leq0.005$ (two-tailed) and corrected for multiple comparisons on the cluster-level at $P\leq0.05$.

Results

Through visual inspection of the event-related averages of the temporal clusters we selected one temporal cluster (ranked 3rd) that showed a speech monitoring effect: Activity for *LISnoise* and *LISvoice* were higher than *PNnoise*, which was itself higher than *PNvoice* (see Fig. 3B; the average temporal correlation between the temporal cluster members was 0.16 ± 0.013). Supplementary Fig. S3 shows the first 10 clusters of temporal components (ranked according to ascending average intra-cluster distance). Quantification of the speech monitoring effect showed a significant main effect for source ($F_{1,11}=9.0$, $P=0.012$) and a significant two-way interaction of source × feedback type ($F_{1,11}=6.2$, $P=0.03$), which supported the observed speech monitoring effect of the temporal components. To ensure that the effect was not due to the choice of using amplitude changes, we repeated the analysis with beta coefficients of the experimental conditions to the time courses. The results of the analysis of beta coefficients corroborated the results of the amplitude changes (see Supplementary results).

Whole-brain random effect analysis using the leave-one-out procedure for building individual design matrices revealed bilateral Heschl's sulcus activity, which was positively related to the leave-one-out model of the temporal component, and left SMA, which was negatively related to the leave-one-out model (see Fig. 3C and Table 2).

Table 2
Leave-one-out random effects results.

	x	y	z	Size	t	p	Area
Pos	51	−21	6	7287	16.6	0.0006	R HS
	−14	−93	5	1952	13.4	0.0005	L OG
	−48	−25	7	6441	10.8	0.0008	L HS
	15	−88	0	2818	9.8	0.0006	R OG
	3	−65	36	1884	6.8	0.0011	R PreC
	−17	−77	−14	418	6.1	0.0010	L OG
	38	−46	39	182	5.7	0.0010	R IPS
	32	−60	40	303	5.1	0.0013	R IPS
	2	−33	−7	227	5.0	0.0012	IC
	44	24	30	143	4.9	0.0014	R MFG
Neg	−1	−10	58	863	−6.2	0.0008	L SMA

Listed are coordinates, size and statistical result of voxel clusters that surpass the cluster-threshold size of 104 mm^3. Pos (Neg), areas showing positive (negative) fit to the temporal component. HS, Heschl's sulcus; OG, occipital gyrus; SMA, supplementary motor area; IPS, intra-parietal sulcus; MFG, middle frontal gyrus; PreC, precuneus; IC, inferior colliculi.

Does the regression of BOLD time courses to the temporal component reveal information about dynamic functional coupling, in the sense that functional connectivity changed with task conditions? To address this question post-hoc, we quantified psychophysiological interactions (PPI) between the regional time courses of areas from the temporal component regression. PPI combines a psychological variable (e.g., task conditions) with a physiological variable (e.g., regional time courses) in order to test if functional connectivity between the physiological variable and an observed time course is augmented by different task conditions (Friston et al., 1997; Kim and Horwitz, 2008). Formally, PPI constitutes a regression analysis onto the psycho- and physiological variables (main effects) and their interaction term, which models the augmentation between the main effects (see Fig. 4A). The psychological variable was obtained by combining the predictors of the original design matrix in such a way that it modeled the differential activity of feedback type of the overt picture naming conditions (i.e., psychological variable = (PNvoice − PNnoise), with all other conditions set to 0), the physiological variable was bilateral Heschl's sulcus (bHS) activity, and the interaction term was modeled as the product of the (mean-centered) main effect variables. This analysis revealed a significant interaction term for SMA across all participants ($t(11) = -3.0$, $P = 0.012$). Fig. 4B shows the scatterplot of SMA activity as a function of bHS activity for a single participant for the PNvoice condition (black dots) and PNnoise condition (open circles). Correlations between bHS and SMA (represented by fitted linear regression lines) changed significantly with type of condition ($r_{PNvoice} = 0.19$, $r_{PNnoise} = 0.33$, $P_{PNvoice - PNnoise} = 0.033$). The PPI interaction terms for the other regions were not significant. These results thus confirm and generalize the finding of a whole-brain network associated with the speech monitoring temporal component, and show dynamic coupling between frontal and temporal regions according to task conditions.

Discussion

We used a sequential combination of spatial and temporal ICA to investigate different brain networks that were related to speech production, perception and monitoring. SICA of the functional data revealed one set of components that showed condition-specific contributions to the different speech tasks, and a second set of components that showed condition-unspecific contributions, which were mostly related to the visual instructions and onset of a task block. TICA of the auditory cortex revealed a single temporal component that represented the speech monitoring pattern. An analysis of the spatial extent of this temporal component showed that bilateral Heschl's sulcus and parietal areas were strongly related to the component, and that SMA was inversely related to the component. Post-hoc psychophysiological interactions further confirmed that task conditions modulated functional connectivity between SMA and bilateral Heschl's sulcus. These findings suggest that speech monitoring is subserved by a dynamic functional coupling between different spatial networks, which changes with task conditions. The results are discussed in detail below.

Condition-specific contribution to speech processes

We found a small set of components that showed condition-specific contributions to overt speech production and comprehension. The AC component comprised areas that are associated with speech

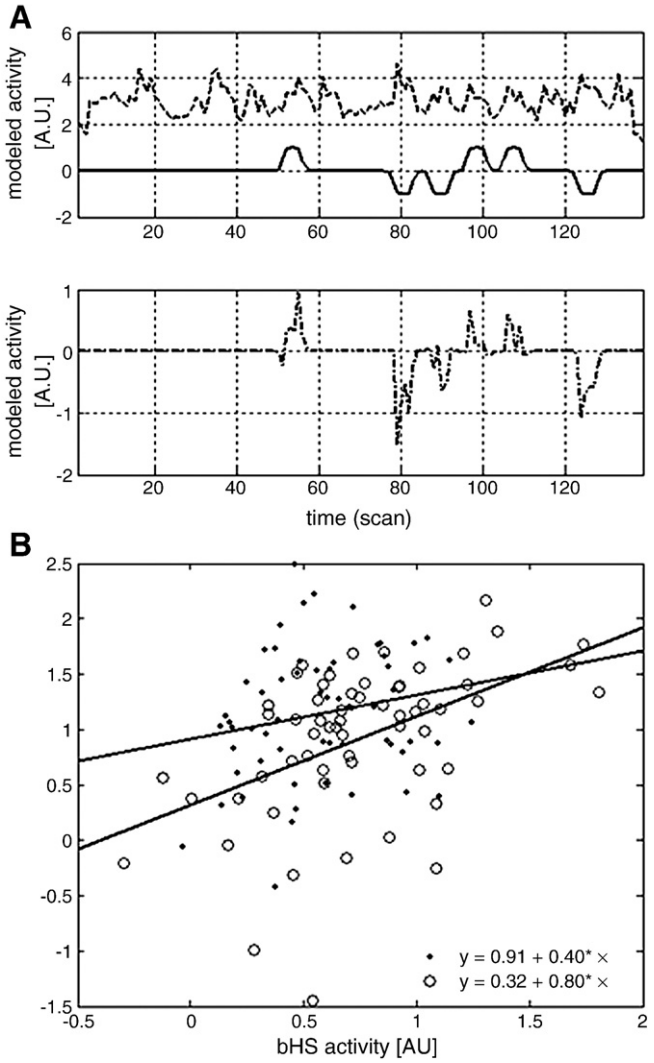

Fig. 4. Psychophysiological interaction results. The figure shows the results of psychophysiological interactions (PPI) between overt picture naming (PN) conditions and bilateral Heschl's sulcus (bHS) activity onto supplementary motor area (SMA) activity. (A) The PPI design matrix for each dataset was modeled in the following way. The psychological variable combined PN conditions of the original (hemodynamically convolved) design matrix such that the PNvoice condition was multiplied by 1 and the PNnoise condition was multiplied by −1 (all other conditions were multiplied by 0). The upper panel (solid line) shows the PPI psychological variable for 1 subject. The physiological variable (regional time course) was obtained by averaging voxel time courses of bHS, as identified by whole-brain correlation with the selected temporal component (bHS activity of 1 subject, broken line). Finally, the PPI interaction term was obtained by multiplying the (mean-centered) psychological and physiological variables (lower panel). (B) Scatterplot of SMA activity as a function of bHS activity for the PNvoice (black dots) and PNnoise (open circles) conditions of 1 subject. Regression lines were fitted to the two data clouds. The change in regression coefficient with different task conditions shows that functional connectivity between SMA and bHS is modulated by task conditions.

perception and production, including primary and secondary auditory cortices, the posterior temporal plane, insula and MTG. The component-activity time course showed a contribution to all speech conditions, but the strongest contribution for overt speech production, whether or not masked with noise. Superior temporal areas have been associated with speech monitoring (Christoffels et al., 2007; McGuire et al., 1996; Toyomura et al., 2007), and have been associated with the production as well as perception of speech (Hickok et al., 2003; Indefrey and Levelt, 2004; Okada and Hickok, 2006; Tzourio-Mazoyer et al., 2004), whereas the insula has been associated with motor aspects of speech production (Ackermann and Riecker, 2004; Dronkers, 1996; Ford et al., 2002) or phonological planning (Bles and Jansma, 2008). The component-activity time course did not clearly show a unique speech monitoring pattern. It should be noted that activity related to the masked speech production condition seemed higher than with normal feedback, but this effect was not significant. Furthermore, this component was more strongly activated for speech production conditions in general, but also to speech comprehension conditions, which suggested that the component represented a mixture of multiple speech processes. Therefore, temporal ICA was used to further disentangle the subcomponent processes in order to investigate if the component was associated with feedback processing in overt speech (see discussion below).

The CS component comprised cortical and subcortical areas that are associated with motor planning and execution, including motor and somatosensory cortex, thalamus and striatum. The time course of this component showed a strong association with overt, but not covert, speech production conditions. Importantly, the CS component also comprised areas of the STG, which in part overlapped with areas found in the AC component. Note that this component specifically reflected speech production processes, because there is no distinction between the two speech production conditions despite the qualitatively quite different sensory inputs, and no response to the two listening control conditions. This finding suggests that sensorimotor processing of overt speech production overlaps with cortical areas of auditory perception, in correspondence with previous studies in humans and monkeys (Foxe et al., 2002; Schroeder et al., 2001).

The left FT component comprised areas in inferior and superior frontal gyri, and temporo-parietal areas, which have been associated with language processing (Binder et al., 1997; Indefrey and Levelt, 2004; Simon et al., 2002; Tzourio-Mazoyer et al., 2004). The component time course showed a significant contribution only to the perception of off-line, prerecorded speech, which suggests that specific frontal and temporo-parietal areas may be uniquely associated with the perception and comprehension of speech, rather than to the production of speech (Hickok et al., 2003; Tzourio-Mazoyer et al., 2004). Albeit speculative, this dissociation between speech comprehension and production may be the consequence of labeling the source of the speech as self vs. externally generated. Alternatively, the left FT component may represent the involvement of attention in language processing. The perception of speech from an external source warrants further processing and attention. Top-down attentional control has been associated with lateral frontal and parietal areas (Corbetta and Shulman, 2002; Hopfinger et al., 2000; Kastner et al., 1999). In addition, attention to speech stimuli can increase activity in left-lateralized temporal areas (Hugdahl et al., 2003).

Finally, the component with strong medial frontal and parietal contributions consistently showed decreased activity from baseline for all speech conditions, which conforms to previous reports of decreased activity during task executions in DMN areas (McKiernan et al., 2003; Raichle et al., 2001). The DMN can be reliably investigated using ICA (Esposito et al., 2006; Greicius et al., 2004; van de Ven et al., 2008). In our study, the DMN showed the strongest deactivations for the speech production tasks, including covert speech production, compared to the listening tasks. This finding is consistent with other fMRI results of decreased activity in DMN areas during off-line perception of one's own speech (Jardri et al., 2007), and extends it to overt speech generation. Furthermore, our results suggest that DMN areas may be modulated by the amount of effort or cognitive control that is required for task performance (Esposito et al., 2006; McKiernan et al., 2003; van de Ven et al., 2008).

Speech monitoring and temporal ICA

None of the spatial components showed a time course that was exclusively related to speech monitoring. In addition, our findings suggested that the AC component represented a temporal mixture of speech processes, including speech production and perception. Reports in the literature further supported this suggestion (Tzourio-Mazoyer et al., 2004).

Using tICA applied to voxels tagged by the spatial AC component, we found one temporal component that exclusively represented the speech monitoring signal in the sense of an interaction of speech source and monitoring difficulty. When estimating the fit of the temporal component to the functional data, results showed the strongest positive fit to bilateral Heschl's sulcus, whereas SMA showed the strongest negative fit. PPI analysis further confirmed that the temporal component represented dynamic coupling between SMA and bilateral Heschl's sulcus. We suggest that these areas represent the speech monitoring network.

The tICA results showed that the speech monitoring network entails a sensory component, which is suppressed in the case of increased monitoring effort, and a motor component, which is increased in the case of increased monitoring effort. These findings contribute to the debate about the information-processing structure of speech monitoring. Computational modeling and neuroimaging studies showed a strong involvement of the motor system in speech monitoring, owing to the role of initiating and controlling motor commands of speech generation (Christoffels et al., 2007; Tourville et al., 2008). Further, other studies suggested that activity of cortical areas of the STG is modulated in accordance with monitoring effort (McGuire et al., 1996). Our results clearly underline that speech monitoring is embedded within (but also crosses) brain networks that contribute to the comprehension or the production of speech. The spatial overlap of speech monitoring areas with other brain networks creates a situation in which the spatial independence criterion of sICA is not optimally suitable to separate speech monitoring into a unique spatial component (cf, Calhoun et al., 2001). The STG appears to be a multimodal processing site, in which sensorimotor and auditory perception processes converge (Eliades and Wang, 2003; Foxe et al., 2002; Schroeder et al., 2001; Tourville et al., 2008), and therefore makes this area a candidate cortical site for speech monitoring.

An unexpected result was the positive correlation between the selected temporal component and bilateral occipital areas. Spatial ICA showed that the visual cortices showed peak of activity 1 or 2 s after block onset, after which activity dropped back to baseline (see Supplementary Fig. S1), and preceded activity peaks of condition-specific clusters. These findings suggest that occipital cortex was mostly responsive to the visual instructions, which preceded block onset by 4 s (i.e., one functional whole-brain image). However, the correlation of early visual areas with the selected temporal component suggests that activity in visual areas may have been modulated during task performance. Application of the tICA and clustering procedures to selected voxels of the extrastriate cortex component (EVC) did not reveal a temporal component cluster similar to the auditory temporal component cluster. Activity in visual areas can be modulated by higher order cognitive processes, such as visual attention or anticipation (Kastner et al., 1999; Tootell et al., 1998), which may have occurred in response to the instructions. Alternatively, activity changes in Heschl's sulcus may have altered activity in visual cortex, which would be in accordance with studies that showed

multisensory interactions of activity between early auditory and visual cortices (Kayser et al., 2008; Lehmann et al., 2006).

General discussion

Our results provide a framework that incorporates several of the different research lines on speech monitoring in a specific manner: Activity in bilateral auditory areas is inversely associated with activity in motor areas, the SMA and bilateral insula. Thus, speech monitoring is not captured within a single, unique spatial network, but comprises the functional coupling between specific nodes of different spatial networks.

These findings may be relevant to a parallel line of research that investigates psychotic symptoms, such as hallucinations in schizophrenia, in which a prevalent neuropsychological model suggests that such symptoms result from an impaired monitoring system of internally generated events (Frith, 1992). A number of studies showed that the on-line perception of auditory verbal hallucinations is associated with increased activity in auditory cortex (Dierks et al., 1999; Lennox et al., 2000; Van de Ven et al., 2005). At the same time, hallucinating schizophrenia patients may be impaired on tasks that include monitoring of self-generated speech or movements (Ford et al., 2002; Johns and McGuire, 1999; Shergill et al., 2005). This impairment has been associated with decreased functional coupling between frontal and temporal areas in hallucinating schizophrenia patients (Ford et al., 2002; Lawrie et al., 2002). These findings, in combination with ours, suggest that hallucinations arise from impaired communication between motor and early auditory areas, resulting in insufficiently suppressed activity in early auditory cortex. In this light, our study provides an experimental design and analysis pipeline that could be applied to overt speech production in schizophrenia in order to further investigate these issues.

The grouping of ICA components in a hierarchical, two-stage fashion is a reliable and intuitive extension of the sogICA framework (Esposito et al., 2005; van de Ven et al., 2008). This methodology preserves the individual decompositions and allows for a hierarchical combination of components in a similar fashion as the nowadays widely used random effects hypothesis test. This study is the first to use the sogICA framework to cluster temporally independent components. The use of tICA in fMRI research is more complex than its spatial variant, but a handful of studies demonstrated that it can be a useful addition to the array of analysis tools in fMRI research (Calhoun et al., 2001), especially in cases where temporal dynamics in the data are not easily modeled using only the experimental protocol (Biswal and Ulmer, 1999; Seifritz et al., 2002).

A potential weakness of our study could be that the line drawing pictures that were used in the picture naming conditions were presented in a scrambled fashion in the listening conditions. This was done in order to prevent participants from automatically engaging in covert picture naming during the listening conditions. Participants may refrain from reflexive covert picture naming during active performance of an alternative task (Jescheniak et al., 2002), but we could not rule out that participants would refrain from covert naming in the case of passive listening. Similarly, we cannot rule out that our choice of presenting visual stimuli may have influenced the results in an unknown way. Nevertheless, our findings conform to previous neurophysiological results and computational modeling predictions, which suggest that the potential influence may be limited.

In conclusion, we demonstrate that different functional networks contribute to different aspects of overt speech production. Speech production and perception networks overlapped in auditory cortex, which suggests that this area is functionally heterogenous. However, speech monitoring was not captured in a single, unique component. Temporal ICA showed that speech monitoring could be captured by a unique temporal component in the auditory cortex. Finally, the temporal component was characterized by a dynamic coupling between frontal and auditory cortices according to task conditions.

Acknowledgments

We kindly acknowledge Prof. Niels Schiller for his general support, Lourens Waldorp for insightful discussions about the analyses and Lourens Waldorp, Bernadette Jansma and anonymous reviewers for providing important comments to previous versions of the manuscript. This study was in part financially supported by grants from the Netherlands Organisation for Scientific Research to VV (grant number 451-07-014) and IKC (grant number 451-06-009).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2009.05.057.

References

Ackermann, H., Riecker, A., 2004. The contribution of the insula to motor aspects of speech production: a review and a hypothesis. Brain Lang. 89, 320–328.
Beer, J.S., 2007. The default self: feeling good or being right? Trends Cogn. Sci. 11, 187–189.
Bell, A.J., Sejnowski, T.J., 1995. An information maximisation approach to blind separation and blind deconvolution. Neural Comput. 7, 1129–1159.
Binder, J.R., Frost, J.A., Hammeke, T.A., Cox, R.W., Rao, S.M., Prieto, T., 1997. Human brain language areas identified by functional magnetic resonance imaging. J. Neurosci. 17, 353–362.
Biswal, B.B., Ulmer, J.L., 1999. Blind source separation of multiple signal sources of fMRI data sets using independent component analysis. J. Comput. Assist. Tomogr. 23, 265–271.
Blakemore, S.J., Rees, G., Frith, C.D., 1998. How do we predict the consequences of our actions? A functional imaging study. Neuropsychologia 36, 521–529.
Bles, M., Jansma, B.M., 2008. Phonological processing of ignored distractor pictures, an fMRI investigation. BMC Neurosci. 9, 20.
Calhoun, V.D., Adali, T., Pearlson, G.D., Pekar, J.J., 2001. Spatial and temporal independent component analysis of functional MRI data containing a pair of task-related waveforms. Hum. Brain Mapp. 13, 43–53.
Christoffels, I.K., Formisano, E., Schiller, N.O., 2007. Neural correlates of verbal feedback processing: an fMRI study employing overt speech. Hum. Brain Mapp. 28, 868–879.
Comon, P., 1994. Independent component analysis — a new concept? Signal Process. 36, 287–314.
Corbetta, M., Shulman, G.L., 2002. Control of goal-directed and stimulus-driven attention in the brain. Nat. Rev. Neurosci. 3, 201–215.
Creutzfeldt, O., Ojemann, G., Lettich, E., 1989. Neuronal activity in the human lateral temporal lobe. II. Responses to the subjects own voice. Exp. Brain Res. 77, 476–489.
Dierks, T., Linden, D.E.J., Jandl, M., Formisano, E., Goebel, R., Lanfermann, H., Singer, W., 1999. Activation of Heschl's gyrus during auditory hallucinations. Neuron 22, 823–835.
Dronkers, N.F., 1996. A new brain region for coordinating speech articulation. Nature 384, 159–161.
Duann, J.-R., Jung, T.-P., Kuo, W.-J., Yeh, T.-C., Makeig, S., Hsieh, J.-C., Sejnowski, T.J., 2002. Single-trial variability in event-related BOLD signals. NeuroImage 15, 823–835.
Eliades, S.J., Wang, X., 2003. Sensory–motor interaction in the primate auditory cortex during self-initiated vocalizations. J. Neurophysiol. 89, 2194–2207.
Eliades, S.J., Wang, X., 2005. Dynamics of auditory–vocal interaction in monkey auditory cortex. Cereb. Cortex 15, 1510–1523.
Esposito, F., Seifritz, E., Formisano, E., Morrone, R., Scarabino, T., Tedeschi, G., Cirillo, S., Goebel, R., Di Salle, F., 2003. Real-time independent component analysis of fMRI time-series. NeuroImage 20, 2209–2224.
Esposito, F., Scarabino, T., Hyvärinen, A., Himberg, J., Formisano, E., Comani, S., Tedeschi, G., Goebel, R., Seifritz, E., Di Salle, F., 2005. Independent component analysis of fMRI group studies by self-organizing clustering. NeuroImage 25, 193–205.
Esposito, F., Bertolino, A., Scarabino, T., Latorre, V., Blasi, G., Popolizio, T., Tedeschi, G., Cirillo, S., Goebel, R., Di Salle, F., 2006. Independent component model of the default-mode brain function: assessing the impact of active thinking. Brain Res. Bull. 70, 263–269.
Ford, J.M., Mathalon, D.H., Whitfield, S., Faustman, W.O., Roth, W.T., 2002. Reduced communication between frontal and temporal lobes during talking in schizophrenia. Biol. Psychiatry 51, 485–492.
Forman, S.D., Cohen, J.D., Fitzgerald, M., Eddy, W.F., Mintun, M.A., Noll, D.C., 1995. Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. Magn. Reson. Med. 33, 636–647.
Foxe, J.J., Wylie, G.R., Martinez, A., Schroeder, C.E., Javitt, D.C., Guilfoyle, D., Ritter, W., Murray, M.M., 2002. Auditory-somatosensory multisensory processing in auditory association cortex: an fMRI study. J. Neurophysiol. 88, 540–543.
Frith, C., 1992. The cognitive neuropsychology of schizophrenia. Psychology Press, Hove, UK.

Friston, K.J., Buechel, C., Fink, G.R., Morris, J., Rolls, E., Dolan, R.J., 1997. Psychophysiological and modulatory interactions in neuroimaging. NeuroImage 6, 218–229.
Fu, C.H., Vythelingum, G.N., Brammer, M.J., Williams, S.C., Amaro Jr., E., Andrew, C.M., Yaguez, L., van Haren, N.E., Matsumoto, K., McGuire, P.K., 2006. An fMRI study of verbal self-monitoring: neural correlates of auditory verbal feedback. Cereb. Cortex 16, 969–977.
Goebel, R., Esposito, F., Formisano, E., 2006. Analysis of functional image analysis contest (FIAC) data with brainvoyager QX: from single-subject to cortically aligned group general linear model analysis and self-organizing group independent component analysis. Hum. Brain Mapp. 27, 392–401.
Greicius, M.D., Srivastava, G., Reiss, A.L., Menon, V., 2004. Default-mode network activity distinguishes Alzheimer's disease from healthy aging: evidence from functional MRI. Proc. Natl. Acad. Sci. U. S. A. 101, 4637–4642.
Hashimoto, Y., Sakai, K.L., 2003. Brain activations during conscious self-monitoring of speech production with delayed auditory feedback: an fMRI study. Hum. Brain Mapp. 20, 22–28.
Hickok, G., Buchsbaum, B., Humphries, C., Muftuler, T., 2003. Auditory–motor interaction revealed by fMRI: speech, music, and working memory in area Spt. J. Cogn. Neurosci. 15, 673–682.
Hopfinger, J.B., Buonocore, M.H., Mangun, G.R., 2000. The neural mechanisms of top-down attentional control. Nat. Neurosci. 3, 284–291.
Hugdahl, K., Thomsen, T., Ersland, L., Rimol, L.M., Niemi, J., 2003. The effects of attention on speech perception: an fMRI study. Brain Lang. 85, 37–48.
Indefrey, P., Levelt, W.J., 2004. The spatial and temporal signatures of word production components. Cognition 92, 101–144.
Jardri, R., Pins, D., Bubrovszky, M., Despretz, P., Pruvo, J.P., Steinling, M., Thomas, P., 2007. Self awareness and speech processing: an fMRI study. NeuroImage 35, 1645–1653.
Jescheniak, J.D., Schriefers, H., Garrett, M.F., Friederici, A.D., 2002. Exploring the activation of semantic and phonological codes during speech planning with event-related brain potentials. J. Cogn. Neurosci. 14, 951–964.
Johns, L.C., McGuire, P.K., 1999. Verbal self-monitoring and auditory hallucinations in schizophrenia. Lancet 353, 469–470.
Kastner, S., Pinsk, M.A., De Weerd, P., Desimone, R., Ungerleider, L.G., 1999. Increased activity in human visual cortex during directed attention in the absence of visual stimulation. Neuron 22, 751–761.
Kayser, C., Petkov, C.I., Logothetis, N.K., 2008. Visual modulation of neurons in auditory cortex. Cereb. Cortex 18, 1560–1574.
Kim, J., Horwitz, B., 2008. Investigating the neural basis for fMRI-based functional connectivity in a blocked design: application to interregional correlations and psycho-physiological interactions. Magn. Reson. Imaging 26, 583–593.
Lane, H., Tranel, B., 1971. The Lombard sign and the role of hearing in speech. J. Speech Hear Res. 14, 677–709.
Lawrie, S.M., Buechel, C., Whalley, H.C., Frith, C.D., Friston, K.J., Johnstone, E.C., 2002. Reduced frontotemporal functional connectivity in schizophrenia associated with auditory hallucinations. Biol. Psychiatry 51, 1008–1011.
Lehmann, C., Herdener, M., Esposito, F., Hubl, D., di Salle, F., Scheffler, K., Bach, D.R., Federspiel, A., Kretz, R., Dierks, T., Seifritz, E., 2006. Differential patterns of multisensory interactions in core and belt areas of human auditory cortex. NeuroImage 31, 294–300.
Lennox, B.R., Park, S.B., Medley, I., Morris, P.G., Jones, P.B., 2000. The functional anatomy of auditory hallucinations in schizophrenia. Psychiatry Res. 100, 13–20.
Levelt, W.J., Roelofs, A., Meyer, A.S., 1999. A theory of lexical access in speech production. Behav. Brain Sci. 22, 1–38 discussion 38–75.
Lindner, M., Hundhammer, T., Ciaramidaro, A., Linden, D.E., Mussweiler, T., 2008. The neural substrates of person comparison — an fMRI study. NeuroImage 40, 963–971.
Makeig, S., Westerfield, M., Jung, T.P., Covington, J., Townsend, J., Sejnowski, T.J., Courchesne, E., 1999. Functionally independent components of the late positive event-related potential during visual spatial attention. J. Neurosci. 19, 2665–2680.
McGuire, P.K., Silbersweig, D.A., Frith, C.D., 1996. Functional neuroanatomy of verbal self-monitoring. Brain 119 (Pt. 3), 907–917.
McKeown, M.J., Jung, T.-P., Makeig, S., Brown, G., Kindermann, S.S., Lee, T.-W., Sejnowski, T.J., 1998. Spatially independent activity patterns in functional MRI data during the Stroop color-naming task. Proc. Natl. Acad. Sci. U. S. A. 95, 803–810.
McKiernan, K.A., Kaufman, J.N., Kucera-Thompson, J., Binder, J.R., 2003. A parametric manipulation of factors affecting task-induced deactivation in functional neuroimaging. J. Cogn. Neurosci. 15, 394–408.
Okada, K., Hickok, G., 2006. Left posterior auditory-related cortices participate both in speech perception and speech production: neural overlap revealed by fMRI. Brain Lang. 98, 112–117.
Paus, T., Perry, D.W., Zatorre, R.J., Worsley, K.J., Evans, A.C., 1996. Modulation of cerebral blood flow in the human auditory cortex during speech: role of motor-to-sensory discharges. Eur. J. Neurosci. 8, 2236–2246.
Postma, A., 2000. Detection of errors during speech production: a review of speech monitoring models. Cognition 77, 97–132.
Postma, A., Kolk, H., 1992. The effects of noise masking and required accuracy on speech errors, disfluencies, and self-repairs. J. Speech Hear. Res. 35, 537–544.
Raichle, M.E., MacLeod, A.M., Snyder, A.Z., Powers, W.J., Gusnard, D.A., Shulman, G.L., 2001. A default mode of brain function. Proc. Natl. Acad. Sci. U. S. A. 98, 676–682.
Schroeder, C.E., Lindsley, R.W., Specht, C., Marcovici, A., Smiley, J.F., Javitt, D.C., 2001. Somatosensory input to auditory association cortex in the macaque monkey. J. Neurophysiol. 85, 1322–1327.
Seifritz, E., Esposito, F., Hennel, F., Mustovic, H., Neuhoff, J.G., Bilecen, D., Tedeschi, G., Scheffler, K., Di Salle, F., 2002. Spatiotemporal pattern of neural processing in the human auditory cortex. Science 297, 1706–1708.
Shergill, S.S., Samson, G., Bays, P.M., Frith, C.D., Wolpert, D.M., 2005. Evidence for sensory prediction deficits in schizophrenia. Am. J. Psychiatry 162, 2384–2386.
Simon, O., Mangin, J.F., Cohen, L., Le Bihan, D., Dehaene, S., 2002. Topographical layout of hand, eye, calculation, and language-related areas in the human parietal lobe. Neuron 33, 475–487.
Talairach, J., Tournoux, P., 1988. Co-planar Stereotaxic Atlas of the Human Brain. Thieme Medical, New York.
Tootell, R.B., Hadjikhani, N., Hall, E.K., Marrett, S., Vanduffel, W., Vaughan, J.T., Dale, A.M., 1998. The retinotopy of visual spatial attention. Neuron 21, 1409–1422.
Tourville, J.A., Reilly, K.J., Guenther, F.H., 2008. Neural mechanisms underlying auditory feedback control of speech. NeuroImage 39, 1429–1443.
Toyomura, A., Koyama, S., Miyamaoto, T., Terao, A., Omori, T., Murohashi, H., Kuriki, S., 2007. Neural correlates of auditory feedback control in human. Neuroscience 146, 499–503.
Tzourio-Mazoyer, N., Josse, G., Crivello, F., Mazoyer, B., 2004. Interindividual variability in the hemispheric organization for speech. NeuroImage 21, 422–435.
Van de Ven, V.G., Formisano, E., Prvulovic, D., Roeder, C.H., Linden, D.E.J., 2004. Functional connectivity as revealed by spatial independent component analysis of fMRI measurements during rest. Hum. Brain Mapp. 22, 165–178.
Van de Ven, V.G., Formisano, E., Röder, C.H., Prvulovic, D., Bittner, R.A., Dietz, M.G., Hubl, D., Dierks, T., Federspiel, A., Esposito, F., Di Salle, F., Jansma, B., Goebel, R., Linden, D. E.J., 2005. The spatiotemporal pattern of auditory cortical responses during verbal hallucinations. NeuroImage 27, 644–655.
van de Ven, V., Bledowski, C., Prvulovic, D., Goebel, R., Formisano, E., Di Salle, F., Linden, D.E., Esposito, F., 2008. Visual target modulation of functional connectivity networks revealed by self-organizing group ICA. Hum. Brain Mapp. 29, 1450–1461.
Wolpert, D.M., Ghahramani, Z., Jordan, M.I., 1995. An internal model for sensorimotor integration. Science 269, 1880–1882.